



Align seg 1/1 to reverse of: AM157357 from: 1 to: 483

```

1 MetSerValProGluGluGluArgLeuProLeuThrGlnArgTr 17
|||||
483 ATGTCGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 434
17 pProArgAlaSerLysPheLeuLeuSerGlyCysAlaIleThrValAlaG 34
|||||
433 GCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 384
34 IuLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGln 50
|||||
383 AGCTACCAACCTTCCCTCGATCTCACAACAACTGACTCCAAATGCCA 334
51 GlyGluAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgGluSerAlaPr 67
|||||
333 GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 284
67 cTyraArgGlyMetValArgThrAlaLeuGlyIleIleGluGluGlyP 84
|||||
283 CATATAGGGGAATGCTGCCAGACAGCCCTAGGATCATTTGAAGAGAGGCT 234
84 heLeuLysLeuTrpGlnGlyValThrProAlaIleThrArgHisValAl 100
|||||
233 TTCTAAAGCTTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 184
101 TyrSerGlyGlyArgMetValThrTyrGlnHisLeuArgGluValAlaPh 117
|||||
183 TATTCGAGAGCTGCAATGCTCATATGAACATCTCCGAGAGAGTGTGT 134
117 eGlyLysSerGluAspGlnHisTyrProLeuTrpLysSerValIleGlyG 134
|||||
133 TGCAGAAAGTGAAGATGACATTAATCCCTTGGAAATCAGTCATGGAG 84
134 IyMetMetAlaGlyValIleGlyGlnPheLeuAlaAsnProThrAspLeu 150
|||||
83 GCATGATGCTGCTGCTTATGGCCAGCTTTTACCAATCCAACTGACCTA 34
151 ValLysValGlnMetGlnMetGluGlyLysArg 161
|||||
33 GTCAAGCTTCAGATGCATGAATGAAGAGAGAGAGAGAGAGAGAGAG 1
seq_name: gb_est42:AM157496
seq_documentation_block:
LOCUS AM157496 483 bp mRNA EST 04-NOV-1999
DEFINITION au79d09.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone
IMAGE:2782481 3' similar to SW:UCP4_HUMAN O95847 MITOCHONDRIAL
UNCOUPLING PROTEIN 4 ; mRNA sequence.
VERSION AM157496
KEYWORDS AM157496
SOURCE AM157496.1 GI:6228897
ORGANISM EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 483)
AUTHORS Hillier,L., Allam,L., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Thelsting,B.,
White,Y., Wyllie,T., Waterston,K. and Wilson,R.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: east@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40UP from Gldco

```

High quality sequence stop: 456.  
Location/Qualifiers

```

1. .483
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2782481"
/clone_lib="Schneider fetal brain 00004"
/sex="male"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript SK (Stratagene);
Site_1: SstI; Site_2: XhoI; double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG-3'
and 3' adaptor sequence:
5'-GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG-3'. The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-trapper method for full-length
enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (LNCIB-Area
Science Park, Trieste, Italy)."
```

BASE COUNT 107 a 140 c 109 g 127 t  
ORIGIN

alignment\_scores:  
Quality: 831.00 Length: 161  
Ratio: 5.161 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-397-342-1 x AM157496/rev ..

Align seg 1/1 to reverse of: AM157496 from: 1 to: 483

```

1 MetSerValProGluGluGluArgLeuProLeuThrGlnArgTr 17
|||||
483 ATGTCGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 434
17 pProArgAlaSerLysPheLeuLeuSerGlyCysAlaIleThrValAlaG 34
|||||
433 GCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 384
34 IuLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGln 50
|||||
383 ACCTAGCAACCTTCCCTCGATCTCACAACAACTGACTCCAAATGCCA 334
51 GlyGluAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgGluSerAlaPr 67
|||||
333 GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 284
67 cTyraArgGlyMetValArgThrAlaLeuGlyIleIleGluGluGlyP 84
|||||
283 CATATAGGGGAATGCTGCCAGACAGCCCTAGGATCATTTGAAGAGAGGCT 234
84 heLeuLysLeuTrpGlnGlyValThrProAlaIleThrArgHisValAl 100
|||||
233 TTCTAAAGCTTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 184
101 TyrSerGlyGlyArgMetValThrTyrGlnHisLeuArgGluValAlaPh 117
|||||
183 TATTCGAGAGCTGCAATGCTCATATGAACATCTCCGAGAGAGTGTGT 134
117 eGlyLysSerGluAspGlnHisTyrProLeuTrpLysSerValIleGlyG 134
|||||
133 TGCAGAAAGTGAAGATGACATTAATCCCTTGGAAATCAGTCATGGAG 84
134 IyMetMetAlaGlyValIleGlyGlnPheLeuAlaAsnProThrAspLeu 150
|||||

```

83 GGATGATGGCTGGTGTATTGGCCAGTTT<sup>1</sup>TAGCCAATCCAAC<sup>2</sup>TGACCTA 34

151 VallysValGlnMetGlnMetGluGlyLysArg 161

33 GTGACGTT CAGATGCCAATGGAAGGAAAAAGA 1

seq\_name: gb\_est42:AW162249

seq\_documentation\_block:

LOCUS	AW162249	584 bp	mRNA	EST	09-NOV-1999
DEFINITION	aut5ell.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone				

IMAGE:2782124 3' similar to SW:UCP4\_HUMAN 095847 MITOCHONDRIAL UNCOUPLING PROTEIN 4 ;, mRNA sequence.

ACCESSION	AW162249
VERSION	AW162249.1
FILE	GI:6301282

KEYWORDS	EST.
SOURCE	human.

ORGANISM	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 584)

**AUTHORS**  
Hiller, L., Allen, M., Bowles, L., Dubaque, T., Gelsel, G., Jost, S.,  
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin

J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE	WashU-NCI human EST Project
JOURNAL	Unpublished (1997)
CONTRIBUTOR	Unpublished (1997)
DATE	1997-11-01
FILE	1997-11-01

COMMENT  
Other\_ESTs: au7sell.y1  
Contact: Wilson RK  
Additional information about the contact:

Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel.: 314 362 1800

Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [cost@unicef.org](mailto:cost@unicef.org)

Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
This clone is available royalty-free through LINT ; contact the  
LINT Consortium ([info@linc.org](mailto:info@linc.org)) for details.

IMAGE Consortium ([infoelimage.lnl.gov](http://infoelimage.lnl.gov)) for further information.  
Seq primer: -40UP from Gibco

High quality sequence stop: 463.  
Location/Qualifiers  
1 584  
FEATURES  
SOURCE

```
source
1. .584
/organism="Homo sapiens"
/ab_xref="taxon:9606"
```

Seq primer: -40RP from Gibco  
High quality sequence stop: 430.  
FEATURES  
Location/Qualifiers  
source

1. 633  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="IMAGE:2782124"  
/clone\_lib="Schneider fetal brain 00004"  
/sex="male"  
/tissue\_type="frontal lobe"  
/dev\_stage="5 months post-conception"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pBluescript SK (Stratagene); Site\_1: Sct1; Site\_2: Xho1; Double-stranded cDNA was prepared from human fetal brain tissue. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence:  
5'-GAGAGAGAGAGAGAGCTCAAGATCCTTAATTAATTAATCCCCCCCC-3' and 3' adaptor sequence:  
5'-GAGAGAGAGAGAGAGCTGATTTTATTTTATTTT-3'. The library was size-selected for >0.5 kb inserts and has an average insert size estimated at 1.2 kb. This library was constructed using the CAP-trapper method for full-length enrichment and has not undergone amplification. Library was constructed by Dr. Claudio Schneider (LNCIB-Area Science Park, Trieste, Italy)."

BASE COUNT 145 a 160 c 198 g 129 t 1 others  
ORIGIN

alignment\_scores:  
Quality: 767.00 Length: 153  
Ratio: 5.079 Gaps: 0  
Percent Similarity: 98.693 Percent Identity: 97.386

alignment\_block:  
US-09-397-342-1 x AM160749 ..

Align seg 1/1 to: AM160749 from: 1 to: 633

```

1 MetServAlProGluGluGluGluArgLeuProLeuThrGlnArgTr 17
175 ATGTCCGTCGCCGAGAGAGAGAGAGCTTTGCGCTGACCCAGAGATG 224
17 pProATgAlaSerLysPheLeuLeuSerGlyCysAlaAlaThrValaIag 34
225 GCCCGAGGAGCAAAATCTCTACTGTCGCGCTGCGCGCTACCGTGGCCG 274
34 luleuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGln 50
275 AGCTAGCAACCTTTCCCTGATCTCACAAAACCTGCACATCCAAATGCCA 324
51 gLyGluAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgLysSerIaPr 67
325 GAGGAGACAGCTCTTGTCTCGTTGGAGAGCGGTGCAGAGAAATCTGCC 374
67 oTyrrArgLysMetValArgThrAlaLeuGlyIleIleGluGluGlyP 84
375 CTTATAGGGGAATGTGTGGCAGACCCCTAGGATCATTTGAAGAGAGGCT 424
84 heLeuLysLeuTPGInGlyValThrProAlaIleTyrrArgHisValaI 100
425 TTCTTAAGCTTTGGCAAGAGTACACCCCATTTACAGACAGCTAGTGTG 474
101 TyrSerGlyLysArgMetValThrTyrrGluHisLeuArgLysValaIaPh 117
475 TATTTCTGGAGGTGCAATGTCACATATGAACATCTCCGAGAGGTTGTGTN 524
117 eGlyLysSerGluAspGluHisTyrrProLeuThrLysSerValIleGlyG 134
525 TGGCAAAAGTGAAGATGAGCATATATCCCTTTGGAATATGATCATTTGGA 574
134 lymetMetAlaGlyValIleGlyGlnPheLeuAlaAsnProThrAspLeu 150

```

```

|||||
575 GATGATGCGCTGTTATTTGGCCAGATTTTACCAATCCACTGACCTTA 624
151 ValLysVal 153
625 ATGAATCTT 633
seq_name: gb_est39:AV652060

```

seq\_documentation\_block:  
LOCUS AV652060 654 bp mRNA EST 07-SEP-2000  
DEFINITION AV652060 GLC Homo sapiens cDNA clone GLCWD06 3', mRNA sequence.  
ACCESSION AV652060  
VERSION AV652060.1 GI:9873074  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Qian,B., Wu,T., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.  
Homo sapiens cDNA clone  
Unpublished (2000)  
CONTACT: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

TITLE  
JOURNAL  
COMMENT

FEATURES  
source  
Location/Qualifiers  
1. 654  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="GLCWD06"  
/clone\_lib="GLC"  
/tissue\_type="Corresponding non cancerous liver tissue"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 213 a 102 c 163 g 171 t 5 others  
ORIGIN

alignment\_scores:  
Quality: 746.00 Length: 142  
Ratio: 5.254 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-397-342-1 x AV652060 ..

Align seg 1/1 to: AV652060 from: 1 to: 654

```

159 GlyLysArgLysLeuGluGlyLysProLeuArgPheArgGlyValHisH 175
1 GGAAGAAAGCAACTGGAGAGGAGAAACCATTTGCGTGTCTACATCA 50
175 sAlaPheAlaLysIleLeuAlaGluGlyGlyIleArgGlyLeuTrpAlaG 192
51 TGCATTTGCAAAAATCTTACCTAAGAGAGCAATACAGGGCTTTGGCGAG 100
192 lYrTrpValProAsnIleGlnArgAlaAlaLeuValAsnMetGlyAspLeu 208
101 GCTGGGACCCATATACAAAGAGCAGCACTGTGATATGAGAGATTTTA 150
209 ThrThrTyrrAspThrValLysHisTyrrLeuValLeuAsnThrProLeuG 225

```



151 ACCACTATGATACAGTGAACACTACTTGCTATTGATACACCACTTGA 200  
 225 uasasnllekethrhisglyleuserSerleucyserglyleuvala 242  
 201 GGACAAATTCATGACCTACGCTTATCAAGTTTATGTTCTGACTGGTAG 250  
 242 laserileleuglythrprolaaspvallelysserargillemetasn 258  
 251 CTCTATTCTGGAACACACCCGATGTCATCAAAAGCAATATATGAT 300  
 259 GlnProArgaspLysGlnGlyArgGlyLeuLeuTyrLysSerSerThrAs 275  
 301 CAACACGAGATAAACAAGAGGAGGACTTTGTATTAATCATCGACGTA 350  
 275 pcysleuileglnalavalglnGlyGlnGlyPheMetSerLeuTyrLysG 292  
 351 CTGCTGATTCAGGCTGTTCAAGGTGAAGGATTCATGATCATATATAAG 400  
 292 lypheleuproserTyrleuargmet 300  
 401 GCTTTTACCATCTTGCTGAGAAATG 426  
 seq\_name: gb\_est42:AW161176  
 seq\_documentation\_block: 564 bp mRNA EST 09-NOV-1999  
 LOCUS AW161176 au79d09.y1 schneider fetal brain 00004 Homo sapiens cDNA clone  
 DEFINITION IMAGE:2782481 5' similar to SW:0CP4\_HUMAN 095847 MITOCHONDRIAL  
 UNCOUPLING PROTEIN 4 ; mRNA sequence.  
 ACCESSION AW161176  
 VERSION AW161176.1 GI:6300209  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 564)  
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S.,  
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Mairi,M., Martin,  
 J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,  
 White,Y., Wylie,T., Waterson,R. and Wilson,R.  
 WASHU-NCI human EST project  
 Unpublished (1997)  
 Other ESTs: au79d09.x1  
 COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estewatson.wustl.edu  
 This clone is available royalty-free through LNC ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 435.  
 Location/Qualifiers  
 FEATURES  
 1..564  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2782481"  
 /clone\_lib="Schneider fetal brain 00004"  
 /sex="male"  
 /tissue\_type="frontal lobe"  
 /dev\_stage="5 months post-conception"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pBluescript SK (Stratagene);  
 Site:1: SstI; Site:2: XhoI; Double-stranded cDNA was  
 prepared from human fetal brain tissue. 5' and 3'  
 adaptors were used in cloning as follows: 5' adaptor  
 sequence:  
 5'-GAAGAGAGAAGAGCTCAAGATCCTTAATTAATATCCCCCCCCC-3'  
 and 3' adaptor sequence:  
 5'-GAGAGAGAGCTCGAGTTTCTTTTCTTTT-3'. The library was  
 size-selected for >0.5 kb inserts and has an average

insert size estimated at 1.2 kb. This library was  
 constructed using the CAP-trapper method for full-length  
 enrichment and has not undergone amplification. Library  
 was constructed by Dr. Claudio Schneider (LNCIB-Area  
 Science Park, Trieste, Italy)."  
 BASE COUNT 129 a 148 c 180 g 107 t  
 ORIGIN  
 alignment\_scores:  
 Quality: 676.00 Length: 130  
 Ratio: 5.200 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-09-397-342-1 x AW161176 ..  
 Align seg 1/1 to: AW161176 from: 1 to: 564  
 1 MetSerValProGlnGluGlnGluArgLeuProLeuThrGlnArgTr 17  
 175 ATGTCCGTCCCGAGAGAGAGAGAGGCTTTGGCGCTGACCCAGAGATG 224  
 17 ProArgLaserLysPheLeuLeuSerGlyCysAlaAlaThrValAlaG 34  
 225 GCCCGAGACGAGCAATTCCTACTGCTCGGCTGCGCGGCTACCGGCGG 274  
 34 luleuAlaThrPheProLeuaspLeuThrLysThrArgLeuGlnMetGln 50  
 275 AGCTGACCAACCTTCCCTCGATCTCACAAAACGCTCACTCCAATGCA 324  
 51 GlyGlnAlaAlaLeuAlaArgLeuGlnLysPglLysAlaArgLysSerAla 67  
 325 GGAGAGAGAGCTCTTGTGCTGCTGGAGAGCGGTGCAAGAGAAATGCCCC 374  
 67 OTyrArgLysMetValArgThrAlaLeuGlyIlelleGlnGluGlnGly 84  
 375 CTATGGGGGATGTCGCCACAGCCCTTAGGATCTTTAAAGAGAGAGCT 424  
 84 heLeuLysLeuTyrPglGlnGlyValThrProAlaIleTyrArgHisValAl 100  
 425 TTCTTAACCTTTGGCAAGAGAGACACCCGCACTTACAGACAGCTAGT 474  
 101 TyrSerGlyGlyArgMetValThrTyrGlnHisLeuAlaGlnValValAlp 117  
 475 TATCTTGAGGTCGATGTCACATATGACATATCCGAGAGGTTGTGT 524  
 117 eGlyLysSerGlnaspGlnHisTyrProLeuTyrLysSer 130  
 525 TGGCAAAAGTGAAGATGACATTAATCCCTTTGGAATCA 564  
 seq\_name: gb\_est42:AW147976  
 seq\_documentation\_block: 803 bp mRNA EST 22-JUN-2000  
 LOCUS AW147976 da01d08.x1 xenopus laevis oocyte xenopus laevis cDNA clone  
 DEFINITION XENOPUS\_SOURCE\_ID:xlnoc001b16 3 similar to SW:0CP4\_HUMAN 095847  
 MITOCHONDRIAL UNCOUPLING PROTEIN 4 ; mRNA sequence.  
 ACCESSION AW147976  
 VERSION AW147976.1 GI:6195872  
 KEYWORDS EST.  
 SOURCE African clawed frog.  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 Xenopodinae; Xenopus.  
 REFERENCE 1 (bases 1 to 803)  
 AUTHORS Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,  
 Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person,  
 B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,  
 Waterson,R. and Wilson,R.  
 WASHU xenopus EST project, 1999  
 TITLE  
 JOURNAL Unpublished (1999)



size-selected for >0.5 kb inserts and has an average insert size estimated at 1.2 kb. This library was constructed using the CAP-trap method for full-length enrichment and has not undergone amplification. Library was constructed by Dr. Claudio Schneider (LNCIB-Area Science Park, Trieste, Italy). " 1 others

BASE COUNT 115 a 141 c 173 g 96 t 1 others

## alignment\_scores:

Quality: 589.00 Length: 115  
Ratio: 5.122 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-397-342-1 x A1878851 ..

Align seg 1/1 to: A1878851 from: 1 to: 526

```
1 MetSerValProGluGluGluArgLeuLeuProLeuThrGlnArgTr 17
|||||
175 ATGTCCCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 224
|||||
17 PProArgAlaSerLysPheLeuLeuSerGlyCysAlaAlaThrValAlaG 34
|||||
225 GCCCGGAGCGAGCAAAATTCCTACTGCTCCGCTCGCGGCTACCTGGCCG 274
|||||
34 LLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGln 50
|||||
275 AGCTAGACACCTTCCCTCGATCTCAAAAACCTCAGCTCAATGCA 324
|||||
51 GLVGLIAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgGlnSerAlaPr 67
|||||
325 GGAGAGACAGCTCTTGGCTGGGTGGAGACGCTGCAAGAGATCTGCCCC 374
|||||
67 OTYrArgGlyMetValArgThrAlaLeuGlyLeuIleGluGluGluGlyP 84
|||||
375 CTATAGGGGATGTGGCGACAGCCCTAGAGGATCATTTGAAGAGAGAGCT 424
|||||
84 hLeuLeuLysLeuTrpGlnGlyValThrProAlaIleTyrArgHisValAl 100
|||||
425 TTCTAAAGCTNTGGCAAGAGATGACACCCGCAATTTACAGACAGTGTG 474
|||||
101 TyrSerGlyGlyArgMetValThrTyrGlnHisLeuArgGluVal 115
|||||
475 TATCTGAGGTGCAATGTGCATATGACATCTCCGAGAGGTG 519
|||||
```

q\_name: gb\_est69:BE061032

## seq\_documentation\_block:

LOCUS BE061032 607 bp mRNA EST 09-JUN-2000  
DEFINITION QV0-BT0041-061099-033-c02 BT0041 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE061032  
VERSION BE061032.1 GI:8405682

## KEYWORDS

EST.

## SOURCE

human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.

## REFERENCE

## AUTHORS

1 (bases 1 to 607)  
Dias Neto, E., Garcia Correia, R., Verjovski-Almeida, S., Briones, M. R.,  
Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,  
Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H.,  
Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare,  
M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and  
Simpson, A. J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-QV0-BT0041-061  
099-033-c02&ct3=1999-10-06&ft4=1)

Seq primer: puc 18 forward  
High quality sequence start: 17  
High quality sequence stop: 596.  
Location/Qualifiers

## FEATURES

source

1. 607

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="BT0041"

/dev\_stage="Adult"

/note="Organ: breast; Vector: puc18; Site.1: Sma1; Site.2:  
Sma1; A mini-library was made by cloning products derived  
from ONESTES PCR (U.S. Letters Patent application No. 196  
'716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

BASE COUNT 187 a 119 c 143 g 158 t

## ORIGIN

## alignment\_scores:

Quality: 583.00 Length: 187  
Ratio: 4.485 Gaps: 4  
Percent Similarity: 69.519 Percent Identity: 66.310

## alignment\_block:

US-09-397-342-1 x BE061032 ..

Align seg 1/1 to: BE061032 from: 1 to: 607

```
95 lLeTyrArgHisVal..... 99
|||||
164 ATACATAGCATATATAGATATAGCAAGTTACTTTTATGATCT 213
|||||
100 .....ValTyrSerGlyGlyArgMetValThrTyrGlnHisLeu 113
|||||
214 TTTTACCAGCTGATTCCTGGAGGTGCAATGTGCATATGCAATATCTCC 263
|||||
113 rGluValAlaPheGlyLysSerGluAspGlnHisTyrProLeuTrpLys 129
|||||
264 GAGAGCTTGTGTTGGCAAAAGTGAAGATGAGCATATCCCTTTGCAAA 313
|||||
130 SerValIleGlyGlyMetMetAlaGlyValIleGlyGlnPheLeuAlaAs 146
|||||
314 TCAGTCATTTGAGGAGGATGATGCTGTGTATTTAGCCAGTTTATAGCAA 363
|||||
146 nProThrAspLeuValLysValGlnMetGlnMetGluGlyLysArgLysL 163
|||||
364 TCCAACTGACCTAGTGAAGGTTCAGATGCCAAATGAAAGAAAGAAAC 413
|||||
163 euGluGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaLys 179
|||||
414 TCGAAGGAAACCATGTGATTT..... 436
|||||
180 lLeuAlaGluGlyGlyIleArgGlyLeuTrpAlaGlyTrpValProAs 196
|||||
436 ..... 436
|||||
196 nIleGlnArgAlaAlaLeuValAsnMetGlyAspLeuThrThrTyrAsp 212
|||||
437 .....AACCATTTATGAT 449
|||||
213 ThrValLysHisTyrLeuValLeuAsnThrProLeuGluAspAsnIleMe 229
```

```

|||||
450 ACAGTGAACACTACTGTTGATTGATACACCACTTGAGGACATATCAAT 499
229 tTtHrHsGLyLeuSerSerLeuCySerGlyLeuValAlaSerIleLeuG 246
500 GACTCAGCGTTTATCAAGTTTATGTTCTGAGCTGAGCTTCTATTTCTG 549
246 tTtHrProAlaAspValIleLySerArgIleMetAsnGlnProArgasp 262
550 GAACACACACCGATGTCATCAAGACGATATATGAAAT...CCACACACC 596
263 LysGlnGly 265
597 AAGGAGGT 605
seq_name: gb_est99:FL11430

```

## seq\_documentation\_block:

```

LOCUS FL11430 341 bp mRNA EST 12-MAR-1995
DEFINITION HSC2MH061 normalized infant brain cDNA Homo sapiens cDNA clone
C-2wh06, mRNA sequence.
ACCESSION FL11430
VERSION FL11430.1 GI:705728
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 341)
REFERENCE
AUTHORS Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
,M.D., Duprat,S., Houngatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
Mitchell,H., Mariage-Samson,R., Plecu,G., Pouliot,Y.,
Sebastien-Kabakchis,C. and Tessier,A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534
Contact: Genethon
GenexPress-Genethon
Genethon Centre de recherche sur le Genome Humain
1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
GenexPress.library_idt: C; GenexPress_sequence_idt: y1c-2wh06
Seq primer: (-21)M13 universal.

```

## FEATURES

```

source
1..341
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="c-2wh06"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site:1: HindIII;
Site:2: NotI; sex=Female; dev stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total brain
; total mRNA was oligo (dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Souares, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"
BASE COUNT 86 a 85 c 96 g 72 t 2 others
ORIGIN

```

```

alignment_scores:
Quality: 567.00 Length: 113
Ratio: 5.108 Gaps: 0
Percent Similarity: 98.230 Percent Identity: 98.230

```

```

alignment_block:
US-09-397-342-1 x FL11430
Align seg 1/1 to: FL11430 from: 1 to: 341

```

```

10 LeuLeuProLeuThrGlnArgTrpProArgAlaSerLysPheLeuLeuSe 26
3 CTTTGGCCCTGACCCAGATGGCCCCCGAGGAGCAAAATTCCTACTGTGC 52
26 rGlyCysAlaAlaThrValAlaGluLeuValAlaThrPheProLeuAspLeuT 43
53 CGGCTGCGGCGCTACCGCTGGCCGAGCTACCAACCTTTCCCTGGATCTTA 102
43 hrLysThrArgLeuGlnMetGlnGlyGluAlaAlaLeuAlaArgLeuGly 59
103 CAAAACCTCGACTCCAAATGCAAGGAGAAAGACAGCTCTGCTCGGTTGGCA 152
60 AspGlyAlaArgGluSerAlaProTyrArgGlyMetValArgThrAlaLe 76
153 GACGGTGCAAGAAATCTGCCCTATAGGGAAATGGTCGACACACCCCT 202
76 uGlyIleIleGluGluGluGlyPheLeuLysLeuTrpGlnGlyValThrP 93
203 AGGATCATTTGAAGAGAGAGGCTTTCTTAAGCTTTGGCAAGGAGTGACAC 252
93 roAlaIleTyrArgHsValValAlaTyrSerGlyGlyArgMetValThrTyr 109
253 CCGCCATTTACAGACAGCATGATATMTCTGAGGTGATGTCACATAT 302
110 GluHsLeuArgGluValValPheGlyLysSerGluasp 122
303 GAACATCTCCGAGAGGTGTGTTGGCAAAAGTGAAAT 341
seq_name: gb_est78:BE772467

```

## seq\_documentation\_block:

```

LOCUS BE772467 407 bp mRNA EST 20-SEP-2000
DEFINITION RC2-FT0126-270600-011-d05 FT0126 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772467
VERSION BE772467.1 GI:10226125
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 407)
REFERENCE
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagal,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.B., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
Sequence tags
Sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?fl=st2=RC2-FT0126-270
600-011-d05&ft3=2000-06-27&ft4=1)
Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 407.
Location/Qualifiers
1..407

```

## FEATURES

```

source
1..407

```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0126"
/dev_stage="Adult"
```

/note="Organ: prostate; tumor: Vector: puc18; Site1: Sma1; Site2: Sma1; A mini-library was made by cloning products derived from ONESRFS PCR (U.S. Letters Patent application No. 196, 716 - Ludwig Institute for Cancer Research) profiles into the puc18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT	113 a	88 c	83 g	123 t
ORIGIN				

```
alignment_scores:
  Quality: 495.00      Length: 95
  Ratio: 5.211         Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.00000
```

ignment\_block: US-09-397-342-1 x BE772467/rev ..

Align seg 1/1 to reverse of: BE772467 from: 1 to: 407

206 G1A<sup>AspLeu</sup>Th<sup>Th</sup>Tyr<sup>Asp</sup>Ph<sup>Arg</sup>Val<sup>Leu</sup>His<sup>Arg</sup>Tyr<sup>Leu</sup>Val<sup>Leu</sup>Asn<sup>Th</sup> 2222  
405 GGAGATTAAACCACTTATGATACAGTGAACACACTACTGTTGATTTGAATAC 3566  
222 r-ProLeuGluAspAsn<sup>Leu</sup>Met<sup>Th</sup>HisGlyLeuSer<sup>Ser</sup>LeuCys<sup>Ser</sup>G 2399  
355 ACACACTTGGAGCAATATCATACACAGGGTTTATCAAGTTTAAGTCTCG 3066  
239 LyeuVal<sup>Ala</sup>Leu<sup>Ser</sup>Leu<sup>Leu</sup>Gly<sup>Thr</sup>Pro<sup>Ala</sup>Asp<sup>Val</sup>Leu<sup>Leu</sup>Ser<sup>Arg</sup> 2555  
305 GACTGGTACTCTTATTCCTGGGAACACCGCGCATGTCATCAAAAGAGA 2566  
256 L<sup>Leu</sup>Met<sup>Asn</sup>Gln<sup>Pro</sup>Arg<sup>Asp</sup>Lys<sup>Gln</sup>Gly<sup>Arg</sup>Gly<sup>Leu</sup>Leu<sup>Tyr</sup>Lys<sup>Se</sup> 2727  
255 ATATATAATCAACACAGAGATTAACACAGGAAGGGGACTTTGTATAATC 2066  
272 rSer<sup>Thr</sup>Asp<sup>Cys</sup>Leu<sup>Leu</sup>Gln<sup>Ala</sup>Val<sup>Gln</sup>Gly<sup>Gln</sup>Gly<sup>Met</sup>Ser<sup>L</sup> 2899  
205 ATGCACTGACCTGCTTGATTCACAGCGTTCCAAGGTGAAGATTCATGATC 1566  
289 eU<sup>Tyr</sup>Lys<sup>Gly</sup>Ph<sup>Leu</sup>Pro<sup>Ser</sup>T<sup>Pro</sup>Leu<sup>Arg</sup>Met 300  
155 TATATAAAGCTTTTACCATCTTGGCGAAGAAG 121

seq\_name: gb\_est50:AW729699

```
seq_documentation_block: .
```

LOCUS	928 bp	EST	16-NOV-2000
DEFINITION	GA_Ea0025M13	Gossypium arboreum 7-10 dpa fiber library	Gossypium
	arboresc CDNA clone GA_Ea0025M13,	arboresc CDNA clone GA_Ea0025M13, mRNA sequence.	

VERSION AM729699.1 GI:7627300

SOURCE *Gossypium arboreum.*

Eukaryota; Viridiplantae; Embry

Malvales; Malvaceae; *Gossypium*.

AUTHORS Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry

**TITLE** An integrated analysis of the genetics, development, and evolution of the cotton fiber

JOURNAL Unpublished (200

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: [rwjmg@clemons.edu](mailto:rwjmg@clemons.edu)  
 High quality sequence stop: 928  
 Location/Qualifiers  
 source 1..928

```

/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="Laxon:29729"
/clone_1b="Ga0025M13"
/clone_1b="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"

```

```

BASE COUNT      223 a      190 c      261 g      254 t
ORIGIN
/note=Vector:pbk-CMV; Site_1: EcoRI; Site_2: XhoI

```

```
alignment_scores:
  quality: 492.50
  ratio: 2.706
  gaps: 5
Percent Similarity: 72.222
Percent Identity: 43.254
```

```
alignment_block:
US-09-397-342-1 x AW729699
```

Align seg 1/1 to: AW729699 from: 1 to: 928

73 ArpThhLleuLgUgUlllellleGluGluGluGluPheLeuLysLeuTrpG1 89  
 74 CgTGTGCGCCCGCGAGTCTCCGTACCAAGCGCTTCAGGCGTGTAA 97  
 48 CgTGTGCGCCCGCGAGTCTCCGTACCAAGCGCTTCAGGCGTGTAA 97  
 89 nglyValThrProAlaIleTyrArgHisValValTyrSerGlyArgm 106  
 98 AGcTTCCTCCCGCGGTATATTAGCAGTCGTGTACACTCCATTCGGA 147  
 106 etValThrTyrGluHisLeuAlaGluValValPheGlyLysSerGluAsp 122  
 148 TCGCGGGATACGGAATTTAAAGAAATTTGGTG.....AGCGCGGT 188  
 123 GluHisTyrProLeuTrpLysSerValIleGlyGlyMetIleAlaGly 139  
 189 GCGCTCTCTTCTGTGCTCTTAACAGCGCTTGAAGTGGCATTTCCGGCG 238  
 139 I11eGlyGluPheLeuAlaAsnProThrAspLeuValLysValGlnMetG 156  
 239 TATTGCTCAGATTGTGTGGCAAGTCCAGCTGACCTGTGTAAAGTGGAGATG 288  
 156 IlnetGluGlyAsnArgLysLeuGluGlyLysProLeuArgPheArgGly 172  
 289 AAGCGGATGGCGGTTTGATTAACAACAGGGCTTCACTCGATTACAAAGGA 338  
 173 ValHisHisAlaPheAlaLysIleLeuAlaGluGlyIleArgGly.L 189  
 339 CCGGTTGATGCTTTTAACAAATATGTAGCATGGAAGGTCTTGGAGGTGC 388  
 189 euTrpAlaGlyTrpValProAsnIleGlnArgAlaIleAlaLeuAlaMet 205  
 389 TTTTGGAAAGGGGTGGCCAAATGTGCAAGAGGGCTTTTGAAGACATG 438  
 206 GlYAspLeuThrTyrIzAspPheValLysHisTyrLeuValIleAsnTh 222  
 439 GGTGACGTACCCGTATATCATCATGCACAAAGCTTTTGTGATTAATATCA 488  
 222 rProLeuGluAspAsnIleMetThrHisGlyLeuSerSerLeuLysSerG 239  
 489 GATATGTGATGAATACATCTATGACACACATGTGGCATCGATCTGTGAG 538  
 239 IlyeuValAlaSerIleLeuGlyTrpProAlaAspValIleLysSerArg 255  
 539 GTCTCTCTGCTCAATAGCTGAGTGTCCGGCTGAGATTTGTGAAAACAGA 568

```

11name: gb|est78:BE773993
256 ILLEKLSNGLNProARgAsPcysInGlyArgGlyLeuLeuTyrTrpLys 272
   ::::::::::::::::::::::::::::::::::::::::::::::::::::
589 ATGATGATCAACGGCGCTAGTAAAGCAAGG..AATGTATGTCACAAAG 635
272 rSerThrAsPCysLeuIleGlnAlaValGlnGlyGlnGlyPheMetSerL 289
   ||||| ||||| ||||| ::::::::::: |||||
636 CTCATATGATGTGTTGGGAAACAGCGTGAATAATCCAAAGTCTAACACCAT 685
289 eutyTrpAsGlyPheLeuProSerThrPLeuArgMetThrProTrpSerMet 305
   ||::||::: ::::: ||||| ||| :::::::
686 TGTGGAAAAGTGTCTTACTACATGCGGCAAGCGCTGGAGCATTTGCCAATTC 735
306 ValPheTrpLeuThrTyrGluLysIleLeuArgLuuMetSerGlyValSerP 322
   ||||| ||::: ::|| ::::||::: |||||
736 TTTTTCGTGGGTGTCACGAGAG...GTGCGCAAAATGTACAGGCGCTTCTC 782
322 ro 322
   ||
783 CT 784

```

```
seq_documentation_block:
```

LOCUS	BE739993	331 bp	EST	20-SEP-2000
DEFINITION	CV1-FT0170-100700-268-f04	FT0170	Homo sapiens	CDNA, mRNA sequence
ACCESSION	BE739993			
VERSION	BE739993.1	GI:10227051		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			

## REFERENCE

TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	PROC. NATL. ACAD. SCI. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G.

FEATURES  
source

BASE COUNT  
ORIGIN

```

alignment_scores:      length: 93
                        quality: 492.00
                        ratio: 5.290
                        gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-397-342-1  x BE773993/rev  ..

```

alignment\_block:  
US-09-397-342-1 x BE773993/rev

Align seg 1/1 to reverse of: BE773993 from: 1 to: 331

100 ValTyrSerG1VQ1YArgMetValThrYrGlnHisLeuAArgValVa 116  
285 GTGATTTCTGAGAGCTGGAATGGTCACATATGACATCTCCGAGAGTTC 236  
116 lPheGlyLysSerGluAspGlnHisTyrProLeuTrpLysSerValIleG 133  
235 GTTGTGCMAAAGGAAGAAGACCATTTATCCCTTTGGAAATCACTCATTC 166  
133 LysGlyMetMetAlaGlyValI1leG1yGlnPheLeuAlaAsnProThrAsp 149  
185 GAGGAGATGATGGCTGGTGTATTATGGCCAGTTTTRAGCCAATCCACACTGAC 136  
150 LeuValLysValGlnMetGlnMetGluGlyLysArgLysLeuGluGlyLys 166  
135 CTAGCTGAGAGTTCAGATGTCMAATGGAGAAAGAAAGAAACCTGGAGGAAA 86  
166 sProLeuAcTyrPheArgG1ValHisIleAlaPheAlaLysIleLeuAlaG 183  
85 ACCATTGCGCATTTTCGTCGTCATCATCATGATCATTTGCCAAAAATCTTACGTC 36  
183 Luc1yGlyL1leArgGlyLeuTrpAlaGly 192  
35 AAGGAGCAATACGAGAGCGCTTTTGGCAGGC 7

seq\_name: gb\_est78:BE772458

LOCUS	BE772458	393 bp	mRNA	EST	20-SEP-2000
DEFINITION	RC2-FT0126-270600-011-a03	FT0126	Homo sapiens	CDNA	mRNA sequence
ACCESSION	BE772458				
VERSION	BE772458.1	GI:10226116			
KEYWORDS	EST.				
SOURCE	human.				

## REFERENCE

TITLE	Shotgun sequencing of the human transcriptome with ORF expressed
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G.

Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
<http://www.ludwig.org.br/scipops/gethtmlid.pl?l1=6t2-RC2-FP0126-270600-011-40383t3-2000-06-27/6t4=1>  
 Seq. primer: puc 18 forward  
 High quality sequence start: 22

High quality sequence stop: 393.  
Location/Qualifiers  
1. 393  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="FT0126"  
/dev\_stage="Adult"  
/note="Organ: prostate, tumor; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
BASE COUNT 109 a 85 c 79 g 120 t  
ORIGIN

alignment\_scores:  
Quality: 485.00 Length: 95  
Ratio: 5.160 Gaps: 0  
Percent Similarity: 98.947 Percent Identity: 98.947

alignment\_block:  
US-09-397-342-1 x BE772458/rev ..  
Align seg 1/1 to reverse of: BE772458 from: 1 to: 393

206 GlyAspLeuThrThrTyrAspThrValIshsTYrLeuValLeuAsnTh 222  
|||||  
391 GGAGATTAAACCACTTATGATCACTGTAACACTTGTGTAATAC 342  
222 rProLeuGluAspAsnIleMetThrHisGlyLeuSerSerLeuCySerg 239  
|||||  
341 ACCACTTGAAGCAATATCATATCACTACGCTTATCAAGTTATATGTTCTG 292  
239 lYleuValAlaSerIleLeuGlyThrProAlaAspValIleYsSerArg 255  
|||||  
291 GACTGTGACTTCTATCTCTGGGAACACCGCATGTCATCAAAAGCAGA 242  
256 lIleMetAsnGlnProArgAspIysGlnGlyArgGlyLeuLeuTyrLysSe 272  
|||||  
241 ATAATGAATCAACTACGAGATTAACAGAGGAGCTTTGTATTAATTC 192  
272 rSerThrAspCysLeuIleGlnAlaValGlnGlyGluGlyPheMetSerL 289  
|||||  
191 ATCGACTGACTGCTTATTCAGGCTTCAAGGTCAAGGATTCAATGAGTC 142  
289 eUTyrlYsGlyPheLeuProSerTrpLeuArgMet 300  
|||||  
141 TATATAAAGCGCTTTTACCATCTTGCGTGAAGATG 107

seq\_name: qb\_esc78:BE772464

seq\_documentation\_block:  
LOCUS BE772464 412 bp mRNA EST 20-SEP-2000  
DEFINITION RC2-FT0126-270600-011-c01 FT0126 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE772464  
VERSION BE772464.1 GI:10226122  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 412)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Negai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
Brunsteht,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
/M.J., Soares,F., Brenland,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags

TITLE

JOURNAL  
MEDLINE  
COMMENT  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=6t2-RC2-FT0126-270  
600-011-c01&t3=2000-06-27&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 28  
High quality sequence stop: 412.  
Location/Qualifiers  
1..412  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="FT0126"  
/dev\_stage="Adult"  
/note="Organ: prostate, tumor; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
BASE COUNT 113 a 90 c 86 g 123 t  
ORIGIN

alignment\_scores:  
Quality: 481.00 Length: 130  
Ratio: 4.859 Gaps: 1  
Percent Similarity: 76.154 Percent Identity: 74.615

alignment\_block:  
US-09-397-342-1 x BE772464/rev ..  
Align seg 1/1 to reverse of: BE772464 from: 1 to: 412

206 GlyAspLeuThrThrTyrAspThrValIshsTYrLeuValLeuAsnTh 222  
|||||  
410 GGAGATTAAACCACTTATGATCACTGTAACACTTGTGTAATAC 361  
222 rProLeuGluAspAsnIleMetThrHisGlyLeuSerSerLeuCySerg 239  
|||||  
360 ACCACTTGAAGCAATATCATATCACTACGCTTATCAAGTTATATGTTCTG 311  
239 lYleuValAlaSerIleLeuGlyThrProAlaAspValIleYsSerArg 255  
|||||  
310 GACTGTGACTTCTATTCCTGGGAACACCGCATGTCACCAAAAGCAGA 261  
256 lIleMetAsnGlnProArgAspIysGlnGlyArgGlyLeuLeuTyrLysSe 272  
|||||  
260 ATAATGAATCAACTACGAGATTAACAGAGGAGCTTTGTATTAATTC 211  
272 rSerThrAspCysLeuIleGlnAlaValGlnGlyGluGlyPheMetSerL 289  
|||||  
210 ATCGACTGACTGCTTATTCAGGCTTCAAGGTCAAGGATTCAATGAGTC 161  
289 eUTyrlYsGlyPheLeuProSerTrpLeuArgMet..... 300  
|||||  
160 TATATAAAGCGCTTTTACCATCTTGCGTGAAGATGATCGTGGATC 111  
300 ..... 300  
110 TCACAGGCCATCCATGAGACCCCGACAGCATTTTCAAGAAGAATGCA 61  
301 .....ThpTrpSerMetValPheTrpLeu 309

60 AGCCTGACCACTTTCACCTTGGGCAAGAGGTTTGGCCTT 21  
:::|||||::: ||| |||

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2001, 16:10:21 ; Search time 1145.32 seconds  
(without alignments)  
7921.668 Million cell updates/sec

Title: US-09-397-342-2  
1039  
Sequence: 1 ccgagccgcgagcccggtat.....cagatatccatcacacgagc 1039

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 9623517 seqs, 4368049070 residues  
Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:

1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_est4: \*  
5: gb\_est5: \*  
6: gb\_est6: \*  
7: gb\_est7: \*  
8: gb\_est8: \*  
9: gb\_est9: \*  
10: gb\_est10: \*  
11: gb\_est11: \*  
12: gb\_est12: \*  
13: gb\_est13: \*  
14: gb\_est14: \*  
15: gb\_est15: \*  
16: gb\_est16: \*  
17: gb\_est17: \*  
18: gb\_est18: \*  
19: gb\_est19: \*  
20: gb\_est20: \*  
21: gb\_est21: \*  
22: gb\_est22: \*  
23: gb\_est23: \*  
24: gb\_est24: \*  
25: gb\_est25: \*  
26: gb\_est26: \*  
27: gb\_est27: \*  
28: gb\_est28: \*  
29: gb\_est29: \*  
30: gb\_est30: \*  
31: gb\_est31: \*  
32: gb\_est32: \*  
33: gb\_est33: \*  
34: gb\_est34: \*  
35: gb\_est35: \*  
36: gb\_est36: \*  
37: gb\_est37: \*  
38: gb\_est38: \*  
39: gb\_est39: \*  
40: gb\_est40: \*  
41: gb\_est41: \*  
42: gb\_est42: \*  
43: gb\_est43: \*

44: em\_esthum10: \*  
45: em\_esthum11: \*  
46: em\_esthum12: \*  
47: em\_esthum13: \*  
48: em\_esthum14: \*  
49: em\_esthum15: \*  
50: em\_esthum16: \*  
51: em\_esthum17: \*  
52: em\_esthum18: \*  
53: em\_esthum19: \*  
54: em\_esthum20: \*  
55: em\_esthum21: \*  
56: em\_esthum22: \*  
57: em\_esthum23: \*  
58: em\_esthum24: \*  
59: em\_esthum25: \*  
60: em\_esthum26: \*  
61: em\_esthum27: \*  
62: em\_esthum28: \*  
63: em\_esthum29: \*  
64: em\_esthum30: \*  
65: em\_esthum31: \*  
66: em\_esthum32: \*  
67: em\_esthum33: \*  
68: em\_esthum34: \*  
69: em\_esthum35: \*  
70: em\_esthum36: \*  
71: em\_esthum37: \*  
72: em\_esthum38: \*  
73: em\_esthum39: \*  
74: em\_esthum40: \*  
75: em\_esthum41: \*  
76: em\_esthum42: \*  
77: em\_esthum43: \*  
78: em\_esthum44: \*  
79: em\_esthum45: \*  
80: em\_esthum46: \*  
81: em\_esthum47: \*  
82: em\_esthum48: \*  
83: em\_esthum49: \*  
84: em\_esthum50: \*  
85: em\_esthum51: \*  
86: em\_esthum52: \*  
87: em\_esthum53: \*  
88: em\_esthum54: \*  
89: em\_esthum55: \*  
90: em\_esthum56: \*  
91: em\_esthum57: \*  
92: em\_esthum58: \*  
93: em\_esthum59: \*  
94: em\_esthum60: \*  
95: em\_esthum61: \*  
96: em\_esthum62: \*  
97: em\_esthum63: \*  
98: em\_esthum64: \*  
99: em\_esthum65: \*  
100: em\_esthum66: \*  
101: em\_esthum67: \*  
102: em\_esthum68: \*  
103: em\_esthum69: \*  
104: em\_esthum70: \*  
105: em\_esthum71: \*  
106: em\_esthum72: \*  
107: em\_esthum73: \*  
108: em\_esthum74: \*  
109: em\_esthum75: \*  
110: em\_esthum76: \*  
111: em\_esthum77: \*  
112: em\_esthum78: \*  
113: em\_esthum79: \*  
114: em\_esthum80: \*  
115: em\_esthum81: \*  
116: em\_esthum82: \*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

```

190: em_gss_pln1:*
191: em_gss_pln2:*
192: em_gss_pro:*
193: em_gss_rdd1:*
194: em_gss_rdd2:*
195: em_gss_rdd3:*
196: em_gss_rdd4:*
197: em_gss_rdd5:*
198: em_gss_vrt1:*
199: em_gss_vrt2:*
200: em_gss_vrt3:*
201: gb_gss1:*
202: gb_gss2:*
203: gb_gss3:*
204: gb_gss4:*
205: gb_gss5:*
206: gb_gss6:*
207: gb_gss7:*
208: gb_gss8:*
209: gb_gss9:*
210: gb_gss10:*
211: gb_gss11:*
212: gb_gss12:*
213: gb_gss13:*
214: gb_gss14:*
215: gb_gss15:*
216: gb_gss16:*
217: gb_gss17:*
218: gb_gss18:*
219: gb_gss19:*
220: gb_gss20:*
221: gb_gss21:*
222: gb_gss22:*
223: gb_gss23:*
224: gb_gss24:*
225: gb_gss25:*
226: gb_gss26:*
227: gb_gss27:*
228: gb_gss28:*
229: gb_gss29:*
230: gb_gss30:*
231: gb_gss31:*
232: gb_gss32:*
233: gb_gss33:*
234: gb_gss34:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
C 1	508.2	48.9	584	AW162249	AW162249 au75e11.x
C 2	482	46.4	483	AW157357	AW157357 au88a06.x
C 3	482	46.4	483	AW157496	AW157496 au79d09.x
4	477.8	46.0	633	AW160749	AW160749 au75e11.y
5	427	41.1	654	AV652060	AV652060 AV652060
6	416.2	40.1	564	AW161176	AW161176 au79d09.y
7	369.2	35.5	526	AI878851	AI878851 au50e02.y
8	334.2	32.2	341	F11430	F11430 HSC2WH061.n
9	287	27.6	407	BE772467	BE772467 RC2-FT012
C 10	284.4	27.4	393	BE772458	BE772458 RC2-FT012
C 11	283.8	27.3	412	BE772464	BE772464 RC2-FT012
C 12	282.8	27.2	331	BE773393	BE773393 QV1-FT017
13	272.4	26.2	409	BE772459	BE772459 RC2-FT012
14	269.4	25.9	391	BE772466	BE772466 RC2-FT012
15	243.6	23.4	607	BE061032	BE061032 QV0-BT004
C 16	239.8	23.1	803	AW147976	AW147976 da01a08.x
C 17	237.8	22.9	590	BF954659	BF954659 MR3-NN021
18	225.6	21.7	375	AW161833	AW161833 au88a06.y







KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 654)  
AUTHORS Qian,B., Wu,T., Huang,Q., Kang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.  
TITLE Homo sapiens CDNA clone  
JOURNAL Unpublished (2000)  
COMMENT Contact: Zengang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@hgc.sh.cn  
This clone is available at CHGC in Shanghai.  
FEATURES  
source  
1..654  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="GLCCWD06"  
/clone\_1ib="GLC"  
/tissue\_type="Corresponding non cancerous liver tissue"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2: XhoI"  
BASE COUNT 213 a 102 c 163 g 171 t 5 others  
ORIGIN  
Query Match 41.1%; Score 427; DB 31; Length 654;  
Best Local Similarity 100.0%; Pred. No. 4e-117;  
Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 514 ggaagaaagaaactggaagaaacacatgcatctgctggtgcatcatcatgca 573  
DB 1 GGAAAAAGAACTGGAGAAACACATTCGATTCGTCGTACATCATGATTCGA 60  
QY 574 aaatcttgcgtgaaggaatcaagaggtcttgaggcagctgggtaccatataca 633  
DB 61 AAAATCTTGTGCTGAAGAGAAATACGAGGCTTTGGGCGAGCTGGGTACCAATATACAA 120  
QY 634 aagaagacacgtgtaataatgagagattaaacacatgatgacagtgaaacactactg 693  
DB 121 AGAGCGACACTGGTGAATATGGAGATTATACCACTTATGATACAGTAACACTACTTG 180  
QY 694 gtaattacacacacacttgagacaaatcatgactacggttcaagttatgtct 753  
DB 181 GATTGAATATACACACTTGAAGCAATATCATGACTACAGGCTTTATCAAGTTATCTTCT 240  
QY 754 ggaactgtagcttctattcttggaacacacagcgatgcatcaaaagcaaatgaaat 813  
DB 241 GCACGTGTGTCTTATTTCTGGGAACACCGAGTCGATCAAAAAGCAGAAATATGAAT 300  
QY 814 caacacagagataaacaaggaagggacttgtataatactgactgactgactgatt 873  
DB 301 CAACACGAGATTAACAGAAAGGAGACTTTGTATTAATATCACTACACTGCTGATT 360  
QY 874 cagagctgtcaagtgaaagatcatgactatataaagctctttacacacttgagct 933  
DB 361 CAGGCTGTTCAGAGTGAAGAGATTCAATGAGTCTATATAAAGGCTTTTACCATCTTGCTG 420  
QY 934 agaatga 940  
DB 421 AGAATGA 427  
RESULT 6

AM161176 564 bp mRNA EST 09-NOV-1999  
LOCUS au79d09.y1  
DEFINITION Schneider fetal brain 00004 Homo sapiens cDNA clone IMAGE:2782481 5' similar to SW:U0P4\_HUMAN 095847 MITOCHONDRIAL UNCOUPLING PROTEIN 4 ; mRNA sequence.  
ACCESSION AM161176  
VERSION AM161176.1 GI:6300209  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 364)  
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Maira,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theisling,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
TITLE WashU-NCI human EST Project  
JOURNAL Unpublished (1997)  
COMMENT Other ESTs: au79d09.x1  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40RP from Gibco  
High quality sequence stop: 435.  
FEATURES  
source  
1..564  
Location/Qualifiers  
1..564  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2782481"  
/clone\_1ib="Schneider fetal brain 00004"  
/sex="Male"  
/tissue\_type="frontal lobe"  
/dev\_stage="5 months post-conception"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pBluescript SK (Stratagene); Site\_1: SbfI; Site\_2: XhoI; Double-stranded cDNA was prepared from human fetal brain tissue. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence:  
5'-GAGAGAGAGAGAGCTCAAGATCCTTAATTAATTAATCCCCCCCCC-3' and 3' adaptor sequence:  
5'-GAGAGAGAGAGCTCGAGTTTCTTTTCTTTT-3'. The library was size-selected for >0.5 kb inserts and has an average insert size estimated at 1.2 kb. This library was constructed using the CAP-trapper method for full-length enrichment and has not undergone amplification. Library was constructed by Dr. Claudio Schneider (LNCIB-Area Science Park, Trieste, Italy)."  
BASE COUNT 129 a 148 c 180 g 107 t  
ORIGIN  
Query Match 40.1%; Score 416.2; DB 111; Length 564;  
Best Local Similarity 99.3%; Pred. No. 6.7e-114;  
Matches 418; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 9 gataccggttaccgctctgctactgctgaatgctccgctccggagagagagagct 68  
DB 144 GGAGTGGCTTATCGTTCGCTACTGCTGATGTCGTCGCCGGAGAGAGAGAGCT 203  
QY 69 ttggcgcgtgacacagagatgagcccgagagcaaatctcactgctgcgcgcgc 128  
DB 204 TTTGGCGCTGAGCCAGAGATGGCCCCGAGCGAGCAAAATTCCTACGTCTCGGCTGGCGCG 263  
QY 129 taccgctggcggagctagcaaccttcccttgatctcaaaaaaactgaactcaatga 188  
DB 264 TACCGTGGCGGAGCTACCACTTTCCCTGGATCTCACAAAACCTGCAGCTCCAAATGCA 323

BASE COUNT	115 a	141 c	173 g	96 t	1 others
ORIGIN					
Query Match	35.5%	Score 369.2;	DB 103;	Length 526;	
Best Local Similarity	98.9%	Pred. No. 8.6e-100;			
Matches 371;	Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;
OY	9	ggatccgcttgcgtctgcgcgtactgctgaatgcctccgcggagagagagagct	68		
Db	144	GGAGGCGCTTATCGTCTTGCCGCTACTGATGTCCTCCGGAGGAGAGAGAGCT	203		
OY	69	tttcgcgtgaaccagagatggccccgagagacaattcctactgtccgctgcgcgc	128		
Db	204	TTTTGCCCTGACCCAGAGATGGCCCCGAGCAGCAAAATTCCTACGTCCGGCTGGCGCGC	263		
OY	129	taccgtgacgagcagcagaaacctttccctgctatctccaaaactgcattcccaatgca	188		
Db	264	TACCGTGGCCGAGCTAGCAACCTTTCCCTGGATCTCAAAAAACCTGACCTCAAAATGCA	323		
OY	189	aggaaagacagctctgtcgtggttgagagacggtcaagagaatctgcgccctataggg	248		
Db	324	AGGAAACACAGCTCTTGCTCGTTGGAGACGGTCAAGAGATGTGCCCTTATAGGGG	383		
OY	249	aatggtgcacacagccctcagatcatcgtgaagaggaagagccttctaaagctttgcaagg	308		
Db	384	AATGCTGGCGCAACAGCCCTAGGGATCATTTGAAGAGGAAGCTTTCTAAGCTNTGGCAAGG	443		
OY	309	agtgacaccgcgcattcacagacacgtatgtatctctgagagtcgaatgtcacatgca	368		
Db	444	AGTGACACCCCGCATTTACAGACAGCTGTGTATCTTGAGAGTGCAGATGTCAACATATGA	503		
OY	369	acatctccgagagagt	383		
Db	504	ACATCTCCGAGAGGT	518		
RESULT	8				
LOCUS	F11430	341 bp	mRNA	EST	12-MAR-1995
DEFINITION	HS22W4061	normalized infant brain cDNA	Homo sapiens	CDNA clone	
ACCESSION	F11430	C-2wh06,	mRNA	sequence.	
VERSION	F11430.1	GI:705728			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 341)				
JOURNAL	Aufrey,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes				
MEDLINE	,M.D., Duprat,S., Houligatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,				
COMMENT	Mitchell,H., Mariage-Samson,R., Pletu,G., Pouliot,Y.,				
	Sebastiani-Kabaktchis,C. and Tessier,A.				
	IMAGE: molecular integration of the analysis of the human genome				
	and its expression				
	C.R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)				
	95277534				
	Contact: Genethon				
	GenexPress-Genethon				
	Genethon Centre de recherche sur le Genome Humain				
	1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE				
	Tel: 33169472800				
	Fax: 33160778698				
	Email: genexpress@genethon.fr				
	Single read.				
	Genexpress_library_idt: G; Genexpress_sequence_idt: y1c-2wh06				
	Seq primer: (-21)M13-universal.				
	Location/Qualifiers				
	1..341				
FEATURES					
SOURCE					

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="c-2wh06"  
/clone\_lib="normalized infant brain cDNA"  
/sex="female"  
/tissue\_type="total brain"

/dev\_stage="3 months old"  
/note="Organ: brain; Vector: latmid BA; Site: 1: HindIII;  
Site: 2: NotI; sex: female; dev\_stage=3 months old;  
isolate=muscular atrophy patient; tissue\_type=total brain  
; total mRNA was oligo (dT) primed and directionally  
cloned 5' -> 3' into the HindIII -> NotI sites of the  
latmid BA vector. Clone library from B.Soures, Psychiatry  
Dept. Columbia University, USA. Normalization method:  
Bento Soares, P.N.A.S in press"

BASE COUNT 86 a 85 c 96 g 72 t 2 others  
ORIGIN

Query Match 32.2%; Score 334.2; DB 152; Length 341;  
Best Local Similarity 98.5%; Pred. No. 2.4e-89;  
Matches 336; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 65 gactttgcgcgtgaccgagagatggcccgagcgagcaaatctcactgtccgctgcg 124  
|||||  
DB 1 GGGTTTGGCGCGTACCGAGATGGCCCCGAGCGAATTCCTACTGTCGGCGTGGC 60  
QY 125 cggctacgcgtgcccgcgagctagcaaccttcccttgatctcacaanaactcgaatcaca 184  
|||||  
DB 61 CGGCTACCGCTGGCCGAGCTACGACACCTTTCCCTGGATCTCAAAAACCTCGACTCCAAA 120  
QY 185 tgcagaagaaagcaagctctctgtcgtgttggaagcgtgcgaagaatctgccccctata 244  
|||||  
DB 121 TCGAAGCAAGAGAGAGCTTGTGCTGGTTGGAGACGCTCAAGATCTGCCCTTATA 180  
QY 245 ggggaatgtgagcaagccctagagatcattgaagaaggaagcttctcaaaagcttgcg 304  
|||||  
DB 181 GGGGAATGTGGCGACAGCGCTGAGATCATGAAGAGAGCTTTCAAAAGCTTTGGC 240  
QY 305 aaggaatgacacccgcacattacagacacgtatgtatctctggaagtcgaatgtacacat 364  
|||||  
DB 241 AAGGATGACACCGCCGACATTTCAGACACGTAGTATNTCTGAGAGTCCAGATGTCACAT 300  
QY 365 atgaacatctcgaagaggtgtgtgttggaagaagtcgaatg 405  
|||||  
DB 301 ATGAACATCTCCGAGAGTGTGTGTGGCAAAAGTGAAGAT 341

RESULT 9  
772467/c

LOCUS BE772467 407 bp mRNA EST 20-SEP-2000  
DEFINITION RC2-FT0126-270600-011-d05 FT0126 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE772467  
VERSION BE772467.1 GI:10226125  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 407)

AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=kt2-RC2-FT0126-270  
600-011-d05&it3=2000-06-27&it4=1)

Seq primer: puc 18 forward  
High quality sequence start: 13  
High quality sequence stop: 407.  
Location/Qualifiers

FEATURES  
source

1. 407  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="FT0126"  
/dev\_stage="Adult"

/note="Organ: prostate tumor; Vector: puc18; Site: 1: SmaI;  
Site: 2: SmaI; A mini-library was made by cloning products  
derived from ORESSES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

BASE COUNT 113 a 88 c 83 g 123 t  
ORIGIN

Query Match 27.6%; Score 287; DB 139; Length 407;  
Best Local Similarity 100.0%; Pred. No. 3.9e-75;  
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 653 tgggaatttaaccactttagatagcagtgaaacacactgtgattgaatcacaccactg 712  
|||||  
DB 407 TGGGAATTTAACCTTTATGATACAGTGAACACACTCTTGATTAATACACCACTTG 348  
QY 713 aggaacatcatgactcaagcttcaagttatgtctgactgtgactgttctatc 772  
|||||  
DB 347 AGGACATATTCATGACTCAGCGTTTATCAAGTTTATGTTGACTGTGACTGTATATTC 288  
QY 773 tgggaacccagcgcgtatgcatcaaaagcagaataatgaatacaacagagataaacaag 832  
|||||  
DB 287 TGGGAACACCAGCGCATGTCATCAAAAGCAATATGATCAACACGAGATTAACAAG 228  
QY 833 gaaggagactttgtataatcatcagactgactgtgttcaagctgtcaagtgtaag 892  
|||||  
DB 227 GAAGGGACTTTTGTAATATCATCAGCTGACTGCTTGAATTCAGGCTGTCAAGGTGAAG 168  
QY 893 gattcatgagctcatataaaggcttttaccacatctgtgctgagaatg 939  
|||||  
DB 167 GATTCAATGAGTCTATATTAAGGCTTTTACACTCTTGCTGGAGAAAG 121

RESULT 10  
BE772458/c

LOCUS BE772458 393 bp mRNA EST 20-SEP-2000  
DEFINITION RC2-FT0126-270600-011-a03 FT0126 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE772458  
VERSION BE772458.1 GI:10226116  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 393)

AUTHORS Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed



JOURNAL MEDLINE  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=RC2-PT0126-270  
600-011-a036t3-2000-06-27&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 22  
High quality sequence stop: 393.  
Location/Qualifiers  
1. 393

FEATURES  
source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="FT0126"  
/dev\_stage="Adult"  
/note="Organ: prostate; tumor; Vector: puc18; Site: 1: Sma1;  
Site: 2: Sma1; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

BASE COUNT 109 a 85 c 79 g 120 t  
ORIGIN

Query Match 27.4%; Score 284.4; DB 139; Length 393;  
Best Local Similarity 99.7%; Pred. No. 2.3e-74;  
Matches 285; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 654 gggaagatttaacacattatgatacagtgaaacactacttgattgaatacaccacttga 713  
DB 392 GGGAGATTTAACCACTTATACAGTGAACACACTCTGTATGATACACACTTGA 333  
QY 714 ggacaataatgactacaggtttataagttatgttcggactggtgactctatct 773  
DB 332 GGACAAATATATGACTCAGCGTTTATCAAGTTTATGTTGAGACTGGTAGCTTATCT 273  
QY 774 gggaacacacagcgatgtlcaaaagcagaataatgaatcaaccacagagataaacaag 833  
DB 272 GGGAAACACAGCGGATGTCTACAAAGCAGAAATGAATCAACTACAGATTAACAAG 213  
QY 834 aaagggaactttgtataatcatcagactgactgctgattccaaggcttcaaggtaag 893  
DB 212 AAGGGGACTTTTGTATTAATCATCGACTGACTGCTTATTCAGGCTGTCAAGGTGAAG 153  
QY 894 attcatgactctataaaggcttttaccatcttgctgaagatg 939  
DB 152 ATTCAATGAGTCTATATTAAGGCTTTTACCACTTGGCTGGAATG 107

RESULT 11  
LOCUS BE772464/c 412 bp mRNA EST 20-SEP-2000  
DEFINITION RC2-FT0126-270600-011-c01 FT0126 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE772464  
VERSION BE772464.1 GI:10226122  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 412)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,

TITLE Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=RC2-PT0126-270  
600-011-c016t3-2000-06-27&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 28  
High quality sequence stop: 412.  
Location/Qualifiers  
1. 412

FEATURES  
source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="FT0126"  
/dev\_stage="Adult"  
/note="Organ: prostate; tumor; Vector: puc18; Site: 1: Sma1;  
Site: 2: Sma1; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

BASE COUNT 113 a 90 c 86 g 123 t  
ORIGIN

Query Match 27.3%; Score 283.8; DB 139; Length 412;  
Best Local Similarity 99.3%; Pred. No. 3.5e-74;  
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 653 ttggagatttaacacattatgatacagtgaaacactacttgattgaatacaccacttg 712  
DB 412 TGGAGATTTAACCACTTATGATACAGTGAACACTCTGTATGATACACACTTG 353  
QY 713 aggaacaataatgactacaggtttataagttatgttcggactggtgactctatct 772  
DB 352 AGGCAATATATGACTCAGCGTTTATCAAGTTTATGTTGAGACTGGTAGCTTATCT 293  
QY 773 tgggaacacacagcgatgtlcaaaagcagaataatgaatcaaccacagagataaacaag 832  
DB 292 TGGGAACACAGCGGATGTCTACAAAGCAGAAATGAATCAACTACAGATTAACAAG 233  
QY 833 gaagggaactttgtataatcatcagactgactgctgattccaaggcttcaaggtaag 892  
DB 232 GAAGGGGACTTTTGTATTAATCATCGACTGACTGCTTATTCAGGCTGTCAAGGTGAAG 173  
QY 893 gattcatgactctataaaggcttttaccatcttgctgaagatg 939  
DB 172 GATTCAATGAGTCTATATTAAGGCTTTTACCACTTGGCTGGAATG 126

RESULT 12  
LOCUS BE773993/c 331 bp mRNA EST 20-SEP-2000  
DEFINITION QY1-FT0170-100700-268-f04 FT0170 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE773993  
VERSION BE773993.1 GI:10227651  
KEYWORDS EST.

SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=et2-qv1-FT0170-100700-268-f04&t3=2000-07-10&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 331.  
Location/Qualifiers

FEATURES  
source  
1. 331  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="FT0170"  
/dev\_stage="Adult"  
/note="Organ: prostate\_tumor; Vector: puc18; Site\_1: Sma1; Site\_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
BASE COUNT 94 a 87 c 52 g 98 t  
ORIGIN

Query Match 27.2%; Score 282.8; DB 139; Length 331;  
Best Local Similarity 99.3%; Pred. No. 6.6e-74;  
Matches 284; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

332 acgtatgtatcttggaagtcgaatggtacataatgaacatccgagagtggttg 391  
|||  
290 ACCCAGTGTATCTGGAGGTGCAATGTCACATATGAACATCTCCGAGAGTGTGTTTG 231  
|||  
332 gcaaaagtgagatgagcatatcccccttggaaatcaagtcattggagagatgatgctg 451  
|||  
230 GCAAAAGTCMAATGACGCTTATCCCTTGGAAATCACTCATTTGGAGCGATGATGGCTG 171  
|||  
452 ggttatctggcaagttttaagccaatccaactgacctagtgaaagtttcagatgcaaatg 511  
|||  
170 GTGTTATTTGGCCAGTTTATAGCCAAATCCACAGTACCTAGTGAAGGTTCAGATGCAAAATG 111  
|||  
512 aaggaagaaagaaacttgaaagaaacattgcatttcgttggttaccatgcatcttg 571  
|||  
110 AAGGAAAAAGAACTGGAAGGAAAAACATTCGAAATTCGTGTGATCATCATGATTTG 51  
|||  
572 caaaactcttaagcgaaggaaggaatacgaaggttcgttgagcagctg 617  
|||  
50 CAAAAATCTTACCTGAAGGAGGAATACGAGGGCTTTGGCAGGCTG 5

RESULT 13  
BE772459

LOCUS BE772459 409 bp mRNA EST 20-SEP-2000  
DEFINITION RC2-FT0126-270600-011-a07 FT0126 Homo sapiens CDNA, mRNA sequence.  
ACCESSION BE772459  
VERSION BE772459.1 GI:10226117  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=et2-RC2-FT0126-270600-011-a07&t3=2000-06-27&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 30  
High quality sequence stop: 409.  
Location/Qualifiers

FEATURES  
source  
1. 409  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="FT0126"  
/dev\_stage="Adult"  
/note="Organ: prostate\_tumor; Vector: puc18; Site\_1: Sma1; Site\_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
BASE COUNT 126 a 81 c 90 g 112 t  
ORIGIN

Query Match 26.2%; Score 272.4; DB 139; Length 409;  
Best Local Similarity 97.6%; Pred. No. 9.4e-71;  
Matches 287; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

647 tgaatatggaagatcttaacacttctgatacagtgaaacactactt-ggtatgaataca 705  
|||  
1 TGAATATGCGAATTAACACATATATGATGACAGTGAACACACTTAACTAGTATGAATACA 60  
|||  
706 ccaattgaggaacataatcaatgacacaggttatacaagttatgcttgagctgagct 765  
|||  
61 CCACCTTGAGAGCAAAATATCATGACTACCGGTTATCAAGTTTATGTTTCGACTGGTAGCT 120  
|||  
766 tctattctgggaacacccagcgatgcatcaaaagcagaataatgaatacaccagagat 825  
|||  
121 TCTATTCTGGGAACACACACCGATGTCATCAAAACAGAAATATGAATCAACACAGAGAT 180  
|||  
826 aacaagaagaaggaagcttctgataaatcatcagctgagctgttgatcaagctgttca 885  
|||  
181 AAACAAGGAAGGAGGACTTTTGTATTAATCATGACTGACTGCTTGATTCAGGCTGTTCAA 240  
|||  
886 gttgaaggaatcatgactatataaagcttttaacacattgctgaggaatg 939

Db 241 GGTAAGATTCATGACTATATTAAGGCTTTTACCACTTGCGTAGAATG 294

|||||

RESULT 14

LOCUS BE772466 391 bp mRNA EST 20-SEP-2000

DEFINITION RC2-PT0126-270600-011-d04 PT0126 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE772466

VERSION BE772466.1 GI:10226124

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.U.

AUTHORS

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-RC2-PT0126-270600-011-d04&t3=2000-06-27&t4=1>) Seq primer: puc 18 forward High quality sequence start: 8 High quality sequence stop: 391. Location/Qualifiers

FEATURES

source

1. 391

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="PT0126"

/dev\_stage="Adult"

/note="Organ: prostate-tumor; Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No.196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 118 a 80 c 87 g 106 t

ORIGIN

Query Match 25.9%; Score 269.4; DB 139; Length 391; Best local Similarity 99.6%; Pred. No. 7.4e-70; Mismatches 1; Indels 0; Gaps 0; Matches 270; Conservative 0;

Db 669 ttatgatacgttaaacactctgtatgataacacacttgagacataatcaatgac 728

|||||

Db 6 TTATGATACAGTGAACACTACTAGGATGATGAATACACACTTGAGGACATATCATGAC 65

|||||

QY 729 tcaagctttatcaagttatgtctctgagctgtagctctatctctggaacacacgcga 788

|||||

Db 66 TCACGGTTTATCAAGTTATGTTCTGACGCTAGAGCTTTATTTCTGGACACACAGCGCA 125

|||||

QY 789 tgcatacaaaagcaagaataatgaatcaacacagagataaacaaggaaggagactttgt 848

|||||

Db 126 TGTCAATCAAAAGCAAGATATGATCAACCAAGATTAACCAAGAGGAGGACTTTGTA 185

|||||

QY 849 taaatcatgactgactgctgttattcagctgttcaagtggaagatcatgagttcata 908

|||||

Db 186 TAAATCATTCAGACTGACTGCTTGGATTCAGGCTTTCAGAGTGAAGATTCATGATCTATA 245

|||||

QY 909 taaagcttttaccatcttgctgtagaatg 939

|||||

Db 246 TAAAGCTTTTACCACTTGCGTAGAATG 276

|||||

RESULT 15

LOCUS BE061032 607 bp mRNA EST 09-JUN-2000

DEFINITION QV0-BT0041-061099-033-c02 BT0041 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE061032

VERSION BE061032.1 GI:8405682

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.U.

AUTHORS

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-QV0-BT0041-061099-033-c02&t3=1999-10-06&t4=1>) Seq primer: puc 18 forward High quality sequence start: 17 High quality sequence stop: 596. Location/Qualifiers

FEATURES

source

1. 607

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="BT0041"

/dev\_stage="Adult"

/note="Organ: breast; Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 187 a 119 c 143 g 158 t

ORIGIN

Query Match 23.4%; Score 243.6; DB 162; Length 607; Best local Similarity 75.9%; Pred. No. 4.8e-62; Mismatches 4; Indels 113; Gaps 1; Matches 369; Conservative 0;

Db 332 acgttagtgattcttgcgagtcgaatggtcatcatatgaactctccgagagtggtgtt 391

|||||

Db 219 ACCCAGGTATTTCTGGAGGCTGCAATGCTCACATATCGACATCTCCGAGAGTGTGTTG 278

|||||

QY 392 gcaaaagtgaagtgagcattatcccttggaaatcagtcattggaggaatgatgctg 451

|||||

```

Db 279 GCAAAAGTGAAGATGAGCATTTATCCCTTTGGAAATCAGTCATGAGGGATGATGGCTG 338
OY 452 GTGTATTGGCCAGTtTTTtagccaaTccaactgacctagtgaaagTtcagatgcaaatg 511
DB 339 GTGTATTGGCCAGTtTTTtagccaaTccaactgacctagtgaaagTtcagatgcaaatg 398
OY 512 aaggaagaaagaaacttgaaagaaacacatTgcgalttcgltgltacatcatgcatltg 571
DB 399 AAGGAAAAAGAACTGGAAAGCAAAACCATTCG----- 431
OY 572 caaaaatcttagctgaaggaagaatacgaaggcttTggcaggtctggtlaaccaatatac 631
DB 432 ----- 431
OY 632 aaagagcagcacTggtgaatatggyagaltTlaacacttatgatacagTgaaacactact 691
DB 432 -----GATTTAACCACTTATGATACAGTGAACACTACT 465
OY 692 tggTattgaatacaccactTgaggaacaataatacTgaactcaeggtTtatcaagTtatgt 751
DB 466 TGGTATTGAATACACCACCTTGAGGACAATATCATGACTCACGCGTTTATCAACTTATGTT 525
OY 752 ctggaactgtagcttctattctcgggaacacacagccgagtgcatacaaaagcagaataatga 811
DB 526 CTGACCTGGTACCTTCTATTTCTGGGACACACGCCGATGTCAATCCAAAGCAGAAATTAATGA 585
OY 812 atcaac 817
DB 586 ATCCAC 591

```

Search completed: April 28, 2001, 17:51:52  
 Job time: 6091 sec



```

|||||
101 ACCTGACCACTTTCCCTGGATCTCACAAAACCTGCATCCAAATGCAA 150
51 GYGLUALAALALEUALAARGLEUGLYASPGLYALAAAGLYUSERALAPR 67
151 GGAGAGACAGCTCTGCTCGTGGAGACGGTGCAGAGGAATCTGCCCC 200
67 CTYARGGLYMETVALARGTHALALEUGLYLEILEUGLUGLUGLYP 84
201 CATATAGGGGAATGTCGCACAGCCCTAGGATCATTTGAAGAGGAAGGCT 250
84 HELLEULYSLEUTRPGINGLYVALTHPRCOALILETYRARGHISVALVAL 100
251 TTCATAAGCTTTGGCAGAGAGTACACCCGCCATTATACAGACCTAGTG 300
101 TYRSERGGLYARGMETVALTHRTYRGUHSLEUARGGLUVALVALPH 117
301 TATTTGGAGGTCGAATGTCACATATGACATCTCCAGAGGTTGTGTT 350
117 EGLYLYSSERGLUASPGLUHISTYRPROLEUTRPLYSSEVALILEGLYG 134
351 TGGCAAAAGTGAAGATGAGCATTTATCCCTTGGAAATCAGTCATTGGAG 400
134 LYMETVALAGLYVALILEGLYGLINPHELEUALASNPROTTHRASPHEU 150
401 GGATATGCTGGTGTATATGGCCAGTTTATGCCAATCCAACTGACCTTA 450
151 VALLYSVALGIMETGLIMETGLUGLYLSARGLYSLEUGLYLYSPR 167
451 GTGAAGGTTCAAGATGCAATGGAAGAAAAGAAAGTGAAGAAAAC 500
167 OLEUARGPHEARGGLYVALHSHISALAPHEALALYSILEUALAGLUG 184
501 ATTTGGATTTGCTGGTGTACATCATGCAATTTGCCAAAATCTTACCTGAAG 550
184 LYGILILEARGGLYLEUTRPRALAGLYTRPRALPROASNILEGINRGLA 200
551 GAGGATACGAGAGGCTTTGGGAGGCTGGTACCCCAATATACAAAGACCA 600
201 ALALEUALASNMETGLYASPLEUTHRTYRASPTRHVALYSHISTY 217
601 GCACGCGTGAATATGAGGATTTAAACCACTATGATACAGTCAAAACACTA 650
217 TLEUVALLEUASNTHRPROLEUGLYASPSNILLEUTHRHISGLYLEUS 234
651 CTGTGATATGAATACACCACTTGAGGACAATATCATGACTCAGGTTTAT 700
234 ERSERLEUCYSERGLYLEUVALALASERILEUGLYTHPRALASP 250
701 CAAGTTATGTTCTGACTGTGACTTCTATTTCTGGAAACACACCCGAT 750
251 VALILEYSERATGILMETASNGINPROARGASPLYSGLINGLYARGL 267
751 GTCATCAAAACAGATATATGATCAACACGAGATTAACAAGAGAGGG 800
267 YLEULEUTRILYSSERSETRHASPYSLEUILEGINALVALINGLYG 284
801 ACTTTTGTATTAATCATGACTGACTGCTGATTCAGGCTGTCAAGGTG 850
284 LUGLYPHEMETSERLEUTRILYSGLYPHELEUPROSETRIPLEUARGMET 300
851 AAGGATTCATGAGTCTATATAAGGCTTTTACCATCTTGGCTGGAATG 900
301 THPROTRPSERMETVALPHETRIPLEUTHRTYRGULYSILEARGGLIME 317
901 ACCCTTGTCATATGCTGCTCTGCTTACTATGAAAAAATCAGAGAGAT 950
317 TSERGLYVALSERPROPE 323
951 GAGTGAAGTCAATCATTT 969
seq_name: gb_ro2:RNO300162

```

```

seq_documentation_block:
LOCUS      RNO300162      2678 bp      mRNA      06-JAN-2001
DEFINITION Rattus norvegicus mRNA for uncoupling protein UCP-4 (ucp-4 gene),
isoform a.
ACCESSION AJ300162
VERSION   AJ300162.1 GI:12055539
KEYWORDS  Isoform a; ucp-4 gene; uncoupling protein UCP-4.
SOURCE    Norway rat.
ORGANISM  Rattus norvegicus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus.
REFERENCE  1 (bases 1 to 2678)
AUTHORS   Alberati-Giani,D., Gatti,S., Rial,E., Danel,F., Hauser,N.,
           Bubendorf,C. and Barfai,T.
TITLE      Three different isoforms of UCP-4 are expressed in rat preoptic
           anterior hypothalamus
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 2678)
AUTHORS   Alberati-Giani,D.
TITLE      Direct Submission
JOURNAL    Submitted (02-NOV-2000) Alberati-Giani D., Pharma Division
           Preclinical Research, CNS, P. Hoffmann-La Roche Ltd.,
           Grenzachstrasse 124, CH 4070 Basel, SWITZERLAND
FEATURES   Location/Qualifiers
            source
              1..2678
                /organism="Rattus norvegicus"
                /strain="Sprague Dawley"
                /db_xref="taxon:10116"
                /sex="male"
                /tissue_type="brain, preoptic anterior hypothalamus"
                /dev_stage="6 weeks"
                /tissue_lib="UNI-ZAP custom cDNA library"
                /country="Switzerland"
                52..1020
                  /gene="ucp-4"
                  52..1020
                    /gene="ucp-4"
                    /function="uncouples respiration in eukaryotic cells"
                    /codon_start=1
                    /product="uncoupling protein UCP-4, isoform a"
                    /protein_id="CAC20898.1"
                    /db_xref="GI:12055540"
                    /translation="MPPEEESLQPLTQWRPRTSKFLSGCAATVAELATFPDLTKT
                    RLOMGEAALAKLGDGAMESAPYRGMMRTALGILOBEGLKLMQVTPATVRYVYSG
                    GRMVTYEHLEREVYFGKSEDEHYPIKMSVYCGMMAGVYIGFLANPTPLVYQOMESKR
                    RLEGRPLREGVNHAHAKILABEGTIGLWAGWIPNORALVNMGDLLTYDVYKHLV
                    LNTALDENIATHLGLSSGLVNSILGTADVILKSRIMQPRDQGRGLLYSSTDVCV
                    IQAVDGEGLSLYKGLFPLSMRLMTFPMVSMVFWLTYEKIRQLDSVSP"
BASE COUNT      683 a      674 c      676 g      645 t
ORIGIN
alignment_scores:
  Quality: 1603.50      Length: 323
  Ratio: 5.074      Gaps: 1
  Percent Similarity: 97.833      Percent Identity: 94.118
alignment_block:
US-09-397-342-1 x RNO300162
Align seg 1/1 to: RNO300162 from: 1 to: 2678
1 MetSerValProGluGluGluGluArgLeuProLeuThrGlnArgTr 17
|||||
52 ATGCCTTCCCTGAGAGAGAGTGC...CTACAGCCACTCAGCAGAGATG 98
17 ProArgLaseRlysPheLeuLeuSerGlyCysAlaIaIaThrValAlag 34
|||||
99 GCCCGGACAGCAAGTCTACTGTCGCGGCTGCCGCCAACCGTGGCG 148
34 LuLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGln 50
|||||

```



257 AGCTAGCAACCTTTCCCTGATCTCACAACCAACCGCAGTCCGATGCAA 306  
51 GlyGluAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgGluSerAla 67  
|||||  
307 GGGAAGCGCCCTTCTGAAGTGGAGATGGTGCATGGAAATCCGCCCCC 356  
67 oTyraArgGlyMetValArgThrAlaLeuGlyIleIleGluGluGly 84  
|||||  
357 TTACAGGGCATGATGCGACGCGCTGGGATGTTCACAGAGGAAGGCT 406  
84 heLeuLysLeuTrpGlnGlyValThrProAlaIleTyrArgHisVal 100  
|||||  
407 TCCTAAAGTTGGGCAAGAGTGAGCCGCCCATTTACAGACCGTAAGT 456  
101 TyrSerGlyGlyArgMetValThrTyrGluHisLeuArgGluVal 117  
|||||  
457 TACTGTGAGAGTCCGATGGTCACTACGAACATCTCCGGGAAGTTGTT 506  
117 eGlyLysSerGluAspGluHisTyrProLeuTrpLysSerValIleGly 134  
|||||  
507 TGGCAAAAGTGAAGATGAACATTATCCCTCGAAGTCCGTCATTGGAG 556  
134 LysMetAlaGlyValIleGlyGlnPheLeuAlaAsnProThrAspLeu 150  
|||||  
557 GGATGATGGCTGCGCGCATCGACAGCTTTTACCCACACCCACTGACCT 606  
151 ValLysValGlnMetGlnMetGluGlyLysArgLysLeuGluGlySer 167  
|||||  
607 GTGAAGGTCCAGATGCCAATGGAAGGAAACGTAGACTGGAAGGAAC 656  
167 oLeuArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAlaGlu 184  
|||||  
657 GTTGAGATTCCGGGAGTACATCATCGTTTCGCAAAATCTTAGCTGAAG 706  
184 LysGlyIleArgGlyLeuTrpAlaGlyTyrValProAsnIleGlnArgAla 200  
|||||  
707 GAGGATCCGTGGGCTTTGGCAGCGTGATACCAATATTCAGAGAGCA 756  
201 AlaLeuValAsnMetGlyAspLeuThrTyrAspThrValLysHisTyr 217  
|||||  
757 GCACGTTGACATGAGGAGATTAAACCACTTACGACACAGTGAACACAT 806  
217 rLeuValLeuAsnThrProLeuGluAspAsnIleMetThrHisGlyLeu 234  
|||||  
807 CCTGGATTGAACACAGCACTTGAAGACAATATCGCCACTCAGCGCTTAT 856  
234 eTyrLeuLysSerGlyLeuValAlaSerIleLeuGlyThrProAlaAsp 250  
|||||  
857 CCAGTTTGTCTGAGCTGCGCTTCTATCTTGGGAACACAGCCGAT 906  
251 ValIleLysSerArgIleMetAsnGlnProArgAspLysGlnGlyArg 267  
|||||  
907 GTCATCAAAAGCCGAAATATGAACCACTCGACACAAACAAGAGAGG 956  
267 yLeuLeuTyrLysSerSerThrAspCysLeuIleGlnAlaValGlnGly 284  
|||||  
957 GCTTTTGTATTAAGTCACTCACTGAGCGGTGATCCAGCGTTCAGAGG 1006  
284 LuGlyPheMetSerLeuTyrLysGlyPheLeuProSerTrpLeuArgMet 300  
|||||  
1007 AAGGCTTCTCGAGCCGTATTAAGCGTTTCTACCCCTCGGCTGAGAAAG 1056  
seq\_name: gb\_r02:RNO300163  
seq\_documentation\_block:  
LOCUS RNO300163 2620 bp mRNA ROD 06-JAN-2001  
DEFINITION Rattus norvegicus mRNA for uncoupling protein UCP-4 (Ucp-4 gene)  
ACCESSION AF300163  
VERSION AJ300163.1 GI:12055541  
KEYWORDS isoform b; Ucp-4 gene; uncoupling protein UCP-4.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
1 (bases 1 to 2620)  
Alberati-Giani,D., Gatti,S., Rial,E., Danel,F., Hauser,N.,  
Bubendorf,C. and Barfai,T.  
TITLE  
Three different isoforms of UCP-4 are expressed in rat preoptic  
anterior hypothalamus  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 2620)  
Alberati-Giani,D.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (02-NOV-2000) Alberati-Giani D., Pharma Division  
Preclinical Research, CNS, F. Hoffmann-La Roche Ltd.,  
Grenzacherstrasse 124, CH 4070 Basel, SWITZERLAND  
FEATURES  
source  
1..2620  
/organism="Rattus norvegicus"  
/strain="Sprague Dawley"  
/db\_xref="taxon:10116"  
/sex="male"  
/issue\_type="brain, preoptic anterior hypothalamus"  
/dev\_stage="6 weeks"  
/tissue\_type="brain, preoptic anterior hypothalamus"  
/country="Switzerland"  
/gene="ucp-4"  
/protein\_id="P10899.1"  
/db\_xref="GI:12055542"  
/translation="MPPPEESLQPLRNWRPGRSKFLSGCAATVAELATPELDLTK  
RLMOGSAALAKLSDGAMESAPRGMRRLGIVQEGFLKMGQVATLPIRYHYVG  
GRATYTHLEHYVFGKSEDEHYPLMKYVIGIMAGVIGOTFLANFTDLVKVOMEGNR  
RLEKPLPRFQVHHAFAKILAEGLRGLMAGWIPNIGRALVNNGLDLYDYKHYLV  
LNLAEEDNIATHGSLGSLVASILCTPADVTKSRIMNPDROGRLLYKSSIDCV  
IOAVQGEFLSLYGFPLSWLRLMEDTLRHDCDLQATLYPLETGVLSDSFYLRAI  
CLVQSHETNNH"  
BASE COUNT 657 a 661 c 679 g 623 t  
ORIGIN  
alignment\_scores:  
Quality: 1480.50 Length: 300  
Ratio: 5.053 Gaps: 1  
Percent Similarity: 97.667 Percent Identity: 94.333  
alignment\_block:  
US-09-397-342-1 x RNO300163 ..  
Align seg 1/1 to: RNO300163 from: 1 to: 2620  
1 MetSerValPProGluGluGluGluArgLeuLeuProLeuThrGlnArgTyr 17  
|||||  
139 ATGCGCTTCCCTGAGAGGAGTGC...CTACAGCCCACTCCACAGAGAT 185  
17 pProArgAlaSerLysPheLeuLeuSerGlyCysAlaAlaThrValAlaG 34  
|||||  
186 GCCCGGACACGACAGTCTCTACTGTCGCGCGCGCGCAACCGTGGCGG 235  
34 LuLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGln 50  
|||||  
236 AGCTAGCAACCTTTCCCTCGATCTCACAACCAACCGAGATCCAGATGCAA 285  
51 GlyGluAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgGluSerAla 67  
|||||  
286 GGGAAGGTGCCCTTGAAGTGGAGATGGTGCATGGAAATCCGCCCCC 335  
67 oTyraArgGlyMetValArgThrAlaLeuGlyIleIleGluGluGly 84  
|||||



```

336 TTACAGGGGCGATGATCCGACGGCCCTGGGATTTGCCAGAGCAAGGCT 385
84 heLeuLysLeuThrPglnglyValThrProAlaIleTyrArgHisValAl 100
386 TCCAAAGTTGTGGCAAGAGAGGACGCCCATTTACAGACACGTAGG 435
101 TyrSerGlyArgMetValThrTyrGlnHisLeuArgGluValAlaPh 117
436 TACTCTGGAGGTCGGATGTCACCTACCAACATCTCCGGGAAGTTGTGT 485
117 eGlyLysSerGluAspGlnHisTyrProLeuThrLysSerValIleGly 134
486 TGGCAAAAGTGAAGATGAACATTATCCCTCGAAGTCCGTCATTGGAG 535
134 LysMetAlaGlyValIleGlyGlnPheLeuAlaAsnProThrAspLeu 150
536 GGATGATGGCGGCGTCATCGACAGTTTAAAGCCACCCACATGACCTC 585
151 ValLysValGlnMetGlnMetGluLysArgLysLeuGluLysPhe 167
586 GTGAAGGTCGAGATGCAATGGAAGGGAACCTGACCTGGAAGGGAACC 635
167 OleuArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAlaGlu 184
636 CTTGAGATTCCGGGGAGTACATCATGCTTGGCAAAATCTTAGCTGAG 685
184 LysGlyIleArgGlyLeuThrPalaGlyTTPValProAsnIleGlnArgAla 200
686 GAGGATACCGTGCGCTTTGGCGAGGCTGATACCAATATTCAGAGACA 735
201 AlaLeuValAsnMetGlyAspLeuThrThrTyrAspThrValLysHisTyr 217
736 GCACCTTGAGACATGGAGATTTAAACCATTAACGACACAGTGAACACTA 785
217 rLeuValLeuAsnThrProLeuGluAspAsnIleMetThrHisGlyLeu 234
786 CCTGATTATGAACACAGCACTTGAGACAATATGCCACTCAGCGCTTAT 835
234 erSerLeuCysSerGlyLeuValAlaSerIleLeuGlyThrProAlaAsp 250
836 CCAGTTGTGTGTGGACTGTGGCTTATTCGTGGAAACACACCGCAT 885
251 ValIleLysSerArgIleMetAsnGlnProArgAspLysGlnIleArgGly 267
886 GTCATCAAAAGCCGAAATATGACCAACCTCGAGCAAAACAGAGAGAG 935
267 yLeuLeuThrLysSerSerThrAspCysLeuIleGlnAlaValGlnGly 284
936 GCTTTTGTATTAAGTATCATCACTGACTGCGTATCCAGGCTGTTCAAGG 985
284 LuGlyPheMetSerLeuThrLysGlyPheLeuProSerThrLeuArgMet 300
986 AAGGCTTCTGAGCCTGTATAAAGGCTTTCAACCTCGCTGCTGAGAAATG 1035

seq_name: qb_pat1:AX061217
seq_documentation_block:
LOCUS AX061217 1726 bp DNA PAT 22-JAN-2001
DEFINITION Sequence 64 from Patent WO0078953.
ACCESSION AX061217
VERSION AX061217.1 GI:12406353
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1726)
REFERENCE
AUTHORS Lal, P., Yang, J., Yue, H., Hillman, J. L., Tang, Y. T., Bandman, O.,
Burford, N., Baughn, M. R., Azimzal, Y., Lu, D. A., Au-Young, J., and
Patterson, C.
Human transport proteins
Patent: WO 0078953-A 64 28-DEC-2000;
Incyte Genomics, Inc. (US)

```

```

FEATURES
    source
        location/Qualifiers
            1..1726
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 552 a 302 c 399 g 473 t
ORIGIN
alignment_scores:
    Quality: 1186.50 Length: 244
    Ratio: 5.181 Gaps: 1
    Percent Similarity: 93.852 Percent Identity: 92.623
alignment_block:
US-09-397-342-1 x AX061217
Align seg 1/1 to: AX061217 from: 1 to: 1726

95 ILeTyrArgHisVal..... 99
|||||:|||||:
397 ATACATTAAGCATATATGATATATAGCCAAAGTTACCTTTTAAATGATCT 446
100 .....ValTyrSerGlyArgMetValThrTyrGlnHisLeu 113
|||||:|||||:
447 TTTTACCACGATGTATTCGAGAGTCGAATGCTCATATATGAACATCTCC 496
113 rGluValAlaPheGlyLysSerGluAspGlnHisTyrProLeuThrLys 129
|||||:|||||:
497 GAGAGGTTGTGTTGGCAAAAGTGAAGATGACATATATCCCTTTGGAAA 546
130 SerValIleGlyLysMetAlaGlyValIleGlyGlnPheLeuAlaAs 146
|||||:|||||:
547 TCAGTATTTGGAGGAGATGATGCTGCTGTTATTTAGGCCATTTTACCAA 596
146 nProThrAspLeuValLysValGlnMetGlnMetGluLysArgLysL 163
597 TCCACATGACCTAGTGAAGGTTCCAGATGCAATGGAAGGAAAAAGCAAC 646
163 euGluGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaLys 179
647 TGGAGGAAACCAATTCGATTCGTGTATCATCATGATTCACAA 696
180 IleLeuAlaGluGlyLysLeuArgLeuThrPalaGlyTTPValProAs 196
697 ATCTTAGCTGAAGAGGAATATGAGAGGCTTGGGAGGCTGGTACCCAA 746
196 nIleGlnArgAlaAlaLeuValAsnMetGlyAspLeuThrThrTyrAspT 213
747 TATACAAAGAGCAGCAGCTGTGTAATATGAGATTTAACCACTTATGATA 796
213 hrValLysHisTyrLeuValLeuAsnThrProLeuGluAspAsnIleMet 229
797 CAGTAAACACTACTTGTGATTAATATACACCACTTGAGGCAATATATCAG 846
220 ThrHisGlyLeuSerSerLeuCysSerGlyLeuValAlaSerIleLeuGly 246
847 ACTCAGGTTTATCAAGTTATGTCTGACTGCTGAGCTTATATCTGGG 896
246 yThrProAlaAspValIleLysSerArgIleMetAsnGlnProArgAspL 263
897 AACACACGCGCATGTCATCAAAAGAGATATGATATCAACACAGATA 946
263 yGlnGlyArgGlyLeuLeuThrLysSerSerThrAspCysLeuIleGln 279
947 AACAGAGAGGAGGACTTTTGTATTAATCATCAGCTGCTTGTGATGAG 996
280 AlaValGlnGlyGluGlyPheMetSerLeuThrLysGlyPheLeuProse 296
997 GCTGTTCAGAGGGAAGATTCATGAGCTATATTAAGGCTTTTATACCAT 1046
296 rThrLeuArgMetThrProTyrSerMetValPheThrPheThrTyrGluL 313
1047 TTGGCTGGAATGACCCCTTGCTCAATGATGTCTTGGCTTACTATATGAAA 1096

```

313 ys11earg1umetSerg1yValSerProphe 323  
|||||  
1097 AATCAGACAGATGATGAGTGCATCATTT 1128

seq\_name: gb\_htg2:AC009216

seq\_documentation\_block:

LOCUS AC009216 169457 bp DNA HTG 03-DEC-1999  
DEFINITION Drosophila melanogaster chromosome X clone BACR38423 (D900) RPCT-98  
38.H.23 map 16D-16F strain Y; cn bw sp, \*\*\* SEQUENCING IN PROGRESS  
\*\*\* 118 unordered pieces.

ACCESSION AC009216 GI:6513905  
VERSION AC009216.8  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE  
1 (bases 1 to 169457)  
Celisner,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,  
Buenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,  
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,  
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,  
Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,  
Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,  
Richards,S., Sethi,H., Svirskas,R.R., Man,K.H., Webster,D.,  
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.

Sequencing of Drosophila melanogaster  
Unpublished  
2 (bases 1 to 169457)  
Celisner,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,  
Buenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,  
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,  
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,  
Kin,E., Lee,B., Lewis,S., Li,P., Lomutan,M.A., Macda,P.,  
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,  
Pfeiffer,B., Poon,L., Segueira,A., Sethi,H., Snir,E.,  
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and  
Rubin,G.M.

Direct Submission  
Submitted (06-AUG-1999) Drosophila Genome Center, Lawrence Berkeley  
Laboratory, MS 64-121, Berkeley, CA 94720, USA  
On Dec 3, 1999 this sequence version replaced gi:6478893.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our sequence  
archive web site (<http://www.fruitfly.org/sequence/>) or send email  
to [hdg@fruitfly.berkeley.edu](mailto:hdg@fruitfly.berkeley.edu). All contigs in this submission meet  
the following cutoffs: length >= 200 bases.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 118 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 489: contig of 489 bp in length  
\* 490 569: gap of unknown length  
\* 570 1126: contig of 557 bp in length  
\* 1127 1206: gap of unknown length  
\* 1207 1561: contig of 355 bp in length  
\* 1562 1641: gap of unknown length  
\* 1642 2169: contig of 528 bp in length  
\* 2170 2249: gap of unknown length  
\* 2250 2833: contig of 584 bp in length  
\* 2834 2913: gap of unknown length  
\* 2914 3802: contig of 889 bp in length  
\* 3803 3882: gap of unknown length  
\* 3883 4828: contig of 946 bp in length  
\* 4829 4908: gap of unknown length  
\* 4909 5573: contig of 665 bp in length

5574 5653: gap of unknown length  
\* 5654 6272: contig of 619 bp in length  
\* 6273 6352: gap of unknown length  
\* 6353 7267: contig of 915 bp in length  
\* 7268 7348: gap of unknown length  
\* 7349 7901: contig of 553 bp in length  
\* 7902 7981: gap of unknown length  
\* 7982 9022: contig of 1042 bp in length  
\* 9023 9103: gap of unknown length  
\* 9104 9692: contig of 590 bp in length  
\* 9693 9772: gap of unknown length  
\* 9773 10394: contig of 622 bp in length  
\* 10395 10474: gap of unknown length  
\* 10475 10877: contig of 403 bp in length  
\* 10878 10957: gap of unknown length  
\* 10958 11712: contig of 755 bp in length  
\* 11713 11792: gap of unknown length  
\* 11793 12174: contig of 382 bp in length  
\* 12175 12254: gap of unknown length  
\* 12255 12898: contig of 644 bp in length  
\* 12899 14013: contig of 1035 bp in length  
\* 14014 14093: gap of unknown length  
\* 14094 15065: contig of 972 bp in length  
\* 15066 15145: gap of unknown length  
\* 15146 15914: contig of 769 bp in length  
\* 15915 15994: gap of unknown length  
\* 15995 16419: contig of 425 bp in length  
\* 16420 16499: gap of unknown length  
\* 16500 17461: contig of 962 bp in length  
\* 17462 17541: gap of unknown length  
\* 17542 18098: contig of 557 bp in length  
\* 18099 18178: gap of unknown length  
\* 18179 18988: contig of 810 bp in length  
\* 18989 19068: gap of unknown length  
\* 19069 20212: contig of 1144 bp in length  
\* 20213 20292: gap of unknown length  
\* 20293 20817: contig of 525 bp in length  
\* 20818 22197: gap of unknown length  
\* 20898 22277: gap of unknown length  
\* 22198 22348: contig of 1161 bp in length  
\* 22349 23518: gap of unknown length  
\* 23519 24889: contig of 1371 bp in length  
\* 24890 24969: gap of unknown length  
\* 24970 25810: contig of 841 bp in length  
\* 25811 25890: gap of unknown length  
\* 25891 27149: contig of 1259 bp in length  
\* 27150 27229: gap of unknown length  
\* 27230 28261: contig of 1032 bp in length  
\* 28262 28341: gap of unknown length  
\* 28342 29712: contig of 1371 bp in length  
\* 29713 29792: gap of unknown length  
\* 29793 31385: contig of 1593 bp in length  
\* 31386 31466: gap of unknown length  
\* 31467 32803: contig of 1338 bp in length  
\* 32804 32883: gap of unknown length  
\* 32884 34706: contig of 1823 bp in length  
\* 34707 34786: gap of unknown length  
\* 34787 36436: contig of 1650 bp in length  
\* 36437 36516: gap of unknown length  
\* 36517 38395: contig of 1879 bp in length  
\* 38396 38475: gap of unknown length  
\* 38476 39602: contig of 1127 bp in length  
\* 39603 39682: gap of unknown length  
\* 39683 41589: contig of 1907 bp in length  
\* 41590 41670: gap of unknown length  
\* 41671 43475: contig of 1806 bp in length  
\* 43476 43555: gap of unknown length  
\* 43556 45795: contig of 2240 bp in length  
\* 45796 45875: gap of unknown length  
\* 45876 48574: contig of 2699 bp in length  
\* 48575 48654: gap of unknown length

```
* 48655 50881: contig of 2227 bp in length
* 50882 50961: gap of unknown length
* 50962 52818: contig of 1857 bp in length
* 52819 52898: gap of unknown length
* 52899 55256: contig of 2358 bp in length
* 55257 55336: gap of unknown length
* 55337 59040: contig of 3704 bp in length
* 59041 59120: gap of unknown length
* 59121 62581: contig of 3461 bp in length
* 62582 62661: gap of unknown length
* 62662 62665: contig of 3604 bp in length
* 62666 66346: gap of unknown length
* 66347 71526: contig of 5181 bp in length
* 71527 71606: gap of unknown length
* 71607 76897: contig of 5291 bp in length
* 76898 76977: gap of unknown length
* 76978 83732: contig of 6755 bp in length
* 83733 83812: gap of unknown length
* 83813 89149: contig of 5337 bp in length
* 89150 89229: gap of unknown length
* 89230 97652: contig of 8423 bp in length
* 97653 97733: gap of unknown length
* 97734 106245: contig of 8512 bp in length
* 106246 106324: gap of unknown length
* 106325 116692: contig of 10368 bp in length
* 116693 116772: gap of unknown length
* 116773 129843: contig of 13071 bp in length
* 129844 129923: gap of unknown length
* 129924 130419: contig of 496 bp in length
* 130420 130499: gap of unknown length
* 130500 131145: contig of 646 bp in length
* 131146 131226: gap of unknown length
* 131227 131840: contig of 615 bp in length
* 131841 131920: gap of unknown length
* 131921 132349: contig of 629 bp in length
* 132350 132629: gap of unknown length
* 132630 133129: contig of 500 bp in length
* 133130 133209: gap of unknown length
* 133210 133458: contig of 249 bp in length
* 133459 133538: gap of unknown length
* 133539 134232: contig of 714 bp in length
* 134233 134332: gap of unknown length
* 134333 134855: contig of 523 bp in length
* 134856 134935: gap of unknown length
* 134936 135396: contig of 461 bp in length
* 135397 135476: gap of unknown length
* 135477 135988: contig of 512 bp in length
* 135989 136068: gap of unknown length
* 136069 136855: contig of 787 bp in length
* 136856 136935: gap of unknown length
* 136936 137763: contig of 828 bp in length
* 137764 137843: gap of unknown length
* 137844 138304: contig of 461 bp in length
* 138305 138384: gap of unknown length
* 138385 138905: contig of 521 bp in length
* 138906 138985: gap of unknown length
* 138986 139626: contig of 641 bp in length
* 139627 139706: gap of unknown length
* 139707 140425: contig of 719 bp in length
* 140426 141245: gap of unknown length
* 141246 141325: contig of 740 bp in length
* 141326 141325: gap of unknown length
```

## alignment\_scores:

```
Quality: 753.50      Length: 323
Ratio: 3.456          Gaps: 4
Percent Similarity: 67.492  Percent Identity: 46.440
```

## alignment\_block:

US-09-397-342-1 x AC009216

Align seg 1/1 to: AC009216 from: 1 to: 169457

```
68 TyrArgGlyMetValArgThrAlaLeuGlyIleIleGlyGlyGlyPhe 84
|||||.....|  |||||.....|
122104 TACCGGCGATGCTGCGCCACCGCCTTCGGATTCGCGTGAAGAGCGC 122153
84 eLeuLysLeuTrpGlnGlyValThrProAlaIleTyrArgHisValValT 101
|||||.....|  |||||.....|
122154 CCGTAAGCTGTGGCAGGCGCTAACCGCGCGCTACCGACACGCTGCT 122203
101 Yr..... 101
122204 ATAGGTGGGCAATCCCAAACTCTCATTTAGCATAGATCAACCGCCA 122253
101 ..... 101
122254 CAGCCATCTGCGCGCAGCAGCTGTAGCATACGCAATTAAGCAGCA 122303
102 ..... 102
122304 TTCTAACCCAACTTCTCCCTGCTCCATGCTCCATCCGAGCGGTGTA 122353
105 rgMetValThrTyrGlnHisLeuArgGluValAlaPheGlyLysSerGlu 121
|||||.....|  |||||.....|
122354 GGATCTGCAGCTACGACCTGATGCGCAAG...GAGTTACACAGAACGCG 122400
122 AspGlnHisTyrProLeuTrpLysSerValIleGlyGlyMetMetAlaG 138
|||||.....|  |||||.....|
122401 ACCCGAGCTGCGCGCTTGGCAAGTCGCGCTGCGCGCTACAGCGCGG 122450
138 YValIleGlyGlnPheLeuAlaAsnProThrAspLeuValLysValGln 155
|||||.....|  |||||.....|
122451 AGCCCTTCCCGCAGTGGCTTGGCTGCGCGCTGACCTGCAAGGTGCAAA 122500
155 etGlnMetGlnGlyLysArgLysLeuGlnGlyLysProLeuArgPheArg 171
|||||.....|  |||||.....|
122501 TCCAAATGAGGCGCGAGCGCTGATGAGGCCACGAGCGGTGTCAG 122550
172 GlyValHisHisAlaPheAlaLysIleLeuAlaGlnGlyGlyIleArgG 188
|||||.....|  |||||.....|
122551 TCCGCTGGCCATGCTTCCGCGCAGTCGTCGCGCGCGGGAATTAAGG 122600
188 YLeuTrpAlaGlyTrpValProAsnIleGlnArgAlaAlaLeuValAsn 205
|||||.....|  |||||.....|
122601 CCTGTGGAAGGCGAGCATCCGGAATGTCACGAGCGCGTGTGTCATC 122650
205 etGlyAspLeuThrThrTyrAspThrValLysHisTyrLeuValLeuAsn 221
|||||.....|  |||||.....|
122651 TGGGCGACCTACACCATACACACATCAAGCCTGATCATGATCGC 122700
222 ThrProLeuGluAspAsnIleMetThrHisGlyLeuSerLeuGlySer 238
|||||.....|  |||||.....|
122701 CTGCAGATGCCCGAGCTGTACACAGTCAGCTGCTGCTGCTGCTGTC 122750
238 rGlyLeuValAlaSerIleLeuGlyThrProAlaAspValIleLysSerA 255
|||||.....|  |||||.....|
122751 CGGATCTGTGCGAGGATCATGAGGCGCAGCGATGTGTGAAGAGCG 122800
255 rGlyMetAsnGlnProArgAspLysGlnGlyArg..... 266
|||||.....|  |||||.....|
122801 GCATCATGACACAGCCACGACGAGATGCGCGTAGTAGACCAATGCG 122850
267 ..... 267
122851 CGCATTTACTATTGGAAAGCTTAATACCTATGTTATTCGCGAGGGC 122900
268 LeuLeuTyrLysSerSerThrAspCysLeuIleGlnAlaValGlnGly 284
|||||.....|  |||||.....|
122901 CTGCTCATTCGCGATCGTGGAGCTGCTGCGACAGAGCGTTTCGAGA 122950
284 uGlyPheMetSerLeuTyrLysGlyPheLeuProSerTrpLeuArgMet 301
|||||.....|  |||||.....|
122951 GGGCTTGTGGCGCTGACAAAGGCTTTCCTGCTGCTGATGCAATGCG 123000
301 hrProTrpSerMetValPheTrpLeuThrTyrGlnLysIleArgGlnMet 317
```

123001 CGCCGTGATGCTACCTTCTGCTTCGTAACAGATCCGACGATG 123050  
318 SerGIYalSerProphie 323  
123051 ATCGAGCCTCCGCGCTAC 123068

seq\_name: gb\_hyt4:AC012162

seq\_documentation\_block:  
LOCUS AC012162 191504 bp DNA HTG 09-FEB-2000  
DEFINITION Drosophila melanogaster chromosome X locus BACR01N10 (D1115)  
Pct-98.01 N.10 map 16D-16F strain y; cn bw sp, \*\*\* SEQUENCING IN  
PROGRES \*\*\* 142 unordered pieces.

ACCESSION AC012162  
VERSION AC012162.9 GI:6957580  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 191504)  
Celiker,S.E., Aghayani,A., Arcalata,T.T., Baxter,E., Blazej,R.G.,  
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,  
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,  
Hindle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,  
Kearney,L., Lee,B., Lewis,S., Li,P., Lomolan,M.A., Mazda,P.,  
Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,  
Richards,S., Sethi,H., Svirskas,R.R., Wan,K.H., Webster,D.,  
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Sequencing of Drosophila melanogaster  
2 (bases 1 to 191504)  
Celiker,S.E., Aghayani,A., Arcalata,T.T., Baxter,E., Blazej,R.G.,  
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,  
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,  
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,  
Kim,E., Lee,B., Lewis,S., Li,P., Lomolan,M.A., Mazda,P.,  
Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,  
Pfeiffer,B., Poon,L., Sequelira,A., Sethi,H., Shif,E.,  
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieren,L.L. and  
Rubin,G.M.

TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (21-Oct-1999) Drosophila Genome Center, Lawrence Berkeley  
Laboratory, MS 64-121, Berkeley, CA 94720, USA  
On Feb 11, 2000 this sequence version replaced gi:6838825.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our sequence  
archive Web site (<http://www.fruitfly.org/sequence/>) or send email  
to [bdg@fruitfly.berkeley.edu](mailto:bdg@fruitfly.berkeley.edu). All contigs in this submission meet  
the following cutoffs: length >= 200 bases.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 142 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 456: contig of 456 bp in length  
457 536: gap of unknown length  
537 1004: contig of 468 bp in length  
1005 1084: gap of unknown length  
1085 1958: contig of 874 bp in length  
1959 2038: gap of unknown length  
2039 2248: contig of 210 bp in length  
2249 2328: gap of unknown length  
2329 2858: contig of 530 bp in length  
2859 2938: gap of unknown length  
2939 3325: contig of 387 bp in length  
3326 3405: gap of unknown length  
3406 4165: contig of 760 bp in length

4166 4245: gap of unknown length  
4246 4716: contig of 471 bp in length  
4717 4796: gap of unknown length  
4797 5373: contig of 577 bp in length  
5374 5453: gap of unknown length  
5454 6050: contig of 597 bp in length  
6051 6130: gap of unknown length  
6131 6720: contig of 590 bp in length  
6721 6800: gap of unknown length  
6801 7461: contig of 661 bp in length  
7462 7542: gap of unknown length  
7543 8104: contig of 563 bp in length  
8105 8184: gap of unknown length  
8185 8505: contig of 321 bp in length  
8506 8585: gap of unknown length  
8586 9227: contig of 642 bp in length  
9228 9307: gap of unknown length  
9308 10091: contig of 784 bp in length  
10092 10171: gap of unknown length  
10172 11099: contig of 928 bp in length  
11079: gap of unknown length  
11179: contig of 568 bp in length  
11747: gap of unknown length  
11827: gap of unknown length  
11828 12767: contig of 940 bp in length  
12768 12847: gap of unknown length  
12848 13398: contig of 551 bp in length  
13399 13478: gap of unknown length  
13479 14007: contig of 529 bp in length  
14008 14087: gap of unknown length  
14088 14867: contig of 780 bp in length  
14868 14947: gap of unknown length  
14948 15562: contig of 615 bp in length  
15563 15642: gap of unknown length  
15643 16859: contig of 1217 bp in length  
16860 16939: gap of unknown length  
16940 17986: contig of 1047 bp in length  
17987 18066: gap of unknown length  
18067 19341: contig of 1275 bp in length  
19342 19421: gap of unknown length  
19422 20926: contig of 1505 bp in length  
20927 21006: gap of unknown length  
21007 22153: contig of 1147 bp in length  
22154 22233: gap of unknown length  
22234 23236: contig of 1003 bp in length  
23237 23316: gap of unknown length  
23317 24607: contig of 1291 bp in length  
24608 24687: gap of unknown length  
24688 25620: contig of 933 bp in length  
25621 25700: gap of unknown length  
25701 26300: contig of 600 bp in length  
26301 26380: gap of unknown length  
26381 27488: contig of 1108 bp in length  
27489 27568: gap of unknown length  
27569 28215: contig of 647 bp in length  
28216 28295: gap of unknown length  
28296 29440: contig of 1145 bp in length  
29441 29520: gap of unknown length  
29521 30587: contig of 1067 bp in length  
30588 30667: gap of unknown length  
30668 32042: contig of 1375 bp in length  
32043 32122: gap of unknown length  
32123 33740: contig of 1618 bp in length  
33741 33820: gap of unknown length  
33821 35375: contig of 1555 bp in length  
35376 35455: gap of unknown length  
35456 36258: contig of 803 bp in length  
36259 36338: gap of unknown length  
36339 38003: contig of 1665 bp in length  
38004 38084: gap of unknown length  
38085 39509: contig of 1426 bp in length  
39510 39589: gap of unknown length  
39590 40719: contig of 1130 bp in length  
40720 40799: gap of unknown length

```

* 40800 42453: contig of 1654 bp in length
* 42454 42533: gap of unknown length
* 42534 43841: contig of 1308 bp in length
* 43842 43921: gap of unknown length
* 43922 45410: contig of 1489 bp in length
* 45411 45490: gap of unknown length
* 45491 46782: contig of 1292 bp in length
* 46783 46862: gap of unknown length
* 46863 48570: contig of 1708 bp in length
* 48571 48650: gap of unknown length
* 48651 51042: contig of 2392 bp in length
* 51043 51122: gap of unknown length
* 51123 52760: contig of 1638 bp in length
* 52761 52840: gap of unknown length
* 52841 54712: contig of 1872 bp in length
* 54713 54792: gap of unknown length
* 54793 56521: contig of 1729 bp in length
* 56522 56601: gap of unknown length
* 56602 58938: contig of 2337 bp in length
* 58939 59018: gap of unknown length
* 59019 60837: contig of 1819 bp in length
* 60838 60917: gap of unknown length
* 60918 62988: contig of 2071 bp in length
* 62989 63068: gap of unknown length
* 63069 65583: contig of 2515 bp in length
* 65584 65663: gap of unknown length
* 65664 68774: contig of 3111 bp in length
* 68775 68854: gap of unknown length
* 68855 71235: contig of 2381 bp in length
* 71236 71315: gap of unknown length
* 71316 74193: contig of 2798 bp in length
* 74194 74913: gap of unknown length
* 74913 77405: contig of 3212 bp in length
* 77406 77485: gap of unknown length
* 77486 80936: contig of 3451 bp in length
* 80937 81016: gap of unknown length
* 81017 83357: contig of 2341 bp in length
* 83358 83437: gap of unknown length
* 83438 87456: contig of 4019 bp in length
* 87457 87536: gap of unknown length
* 87537 91532: contig of 3996 bp in length
* 91533 91612: gap of unknown length
* 91613 97143: contig of 5531 bp in length
* 97144 97223: gap of unknown length
* 97224 105593: contig of 8370 bp in length
* 105594 105673: gap of unknown length
* 105674 112071: contig of 6398 bp in length
* 112072 112151: gap of unknown length
* 112152 121011: contig of 8860 bp in length
* 121012 121091: gap of unknown length
* 121092 143113: contig of 22022 bp in length
* 143114 143193: gap of unknown length
* 143194 143587: contig of 394 bp in length
* 143588 143667: gap of unknown length
* 143668 143927: contig of 260 bp in length
* 143928 144007: gap of unknown length
* 144008 144842: contig of 835 bp in length
* 144843 144922: gap of unknown length
* 144923 145360: contig of 438 bp in length
* 145361 145440: gap of unknown length
* 145441 145915: contig of 475 bp in length
* 145916 145995: gap of unknown length
* 145996 146532: contig of 537 bp in length
* 146533 146612: gap of unknown length

```

alignment\_scores:      Quality: 753.50      Length: 323  
 Ratio: 3.456      Gaps: 4  
 Percent Similarity: 67.492      Percent Identity: 46.440

alignment\_block:  
 US-09-397-342-1 x AC012162/rev

```

Align seg 1/1 to reverse of: AC012162 from: 1 to: 191504
68 TyrArgLysMetValArgThrAlaLeuGlyIleIleGluGluGlyPhe 84
|||||
136550 TACCGCGCATGTGGCCACCGCCTTGGGATGGCGGTAGAGAGGCC 136301
84 eLeuLysIleuTrpGlnGlyValThrProAlaIleTyrArgHisValVal 101
|||||
136300 CTGAAGCTGTGGCAGGGCGCTAAGCGGGCGCTTACCGACAGCTGTCT 136251
101 yr.....
136250 ATAGTGGCGATTCCCAAACTCCTCATTTAGCATAGATCAGCGCCA 136201
101 .....
136200 CAGCCATCTGCCCGACAGCTGTAGCCATTCAGCATTAAGTCAGCAG 136151
102 .....
136150 TTCTAACCCAACTTCTCCCTCTCCATGTCATCCGACGGCGTGTCA 136101
105 rGmeValThrTyrGluHisLeuArgGlyValValPheGlyLysSerGlu 121
|||||
136100 GGATCTGCACCTAGCAGCTGATGGCAG...GAGTTCACAGAAAGCGC 136054
122 AspGluHisTyrProLeuTrpLysSerValIleGlyLysMetLeuAla 138
|||||
136053 ACCCAGGCTCTCGCGTTGGAAAGTCGGCTGTGGCGCTCACAGCGCG 136004
138 yAlIleGlyInPheLeuAlaAsnProThrAspLeuValValGln 155
|||||
136003 AGCGGTGCCAGAGGCGTTCGCTCGCGCTGACCTGTCAAGTGCAGAA 135954
155 etGlnMetGluGlyLysArgLysLeuGluGlyLysProLeuArgPheArg 171
|||||
135953 TCCAAATGAGAGGCCGACGAGCTGTATGGCGAGCCACGAGGAGTCAC 135904
172 GlyAlaHisHisAlaPheAlaLysIleLeuAlaGluGlyIleArgGlu 188
|||||
135903 TCCGCTGGCATGCTTCGCCAGATGTCAGCGCGGAGTAAGGG 135854
188 yLeuTrpAlaGlyTrpValProAsnIleGlnArgAlaAlaLeuValAsn 205
|||||
135853 CCTGTGAAGGAGCGAGCATCCCGATGTGCAGCAGCGCGCTGTGCATC 135804
205 etGlyAspLeuThrThrTyrAspThrValLysHisTyrLeuValLeuAsn 221
|||||
135803 TGGCGACCTAACCCACATACGACACATCAAGCAGCTGATCATGAATCGC 135754
222 ThrProLeuGluAspAsnIleMetThrHisGlyLeuSerSerLeuGly 238
|||||
135753 CTGCAGATGCCCGCAGCTGTCAACAGTGCAGCTGTGCTCCGTTGTGC 135704
238 rGlyLeuValAlaSerIleLeuGlyThrProAlaAspValIleLysSerA 255
|||||
135703 CGGATTCGTGGCAGCATTCATGGCAGCCAGCTGATGTGTGAAGCGC 135654
255 rGlyIleMetAsnGlnProArgAspLysGlnGlyArg..... 266
|||||
135653 GCATCATGAACACACCGCCAGCAGAGATGGCGGTAAAGTACCAATGCG 135604
267 .....
135603 CGCATTTAATTGGAACGCTTAATCAGCTAATGTTCTCCAGGGGC 135554
268 LeuLeuTyrLysSerSerThrAspCysLeuIleGlnAlaValGlnGly 284
|||||
135553 CTGCTATATCGCGGATCCGAGATGCTCGACACAGCGGTTTCGAAGCA 135504
284 uGlyPheMetSerLeuTyrLysGlyPheLeuProSerTrpLeuArgMet 301
|||||
135503 GGGCTTGTGGCGCTGTACAAAGGCTTTCGCTGTGATACGAATG 135454

```

```

301 hpfctRsrMwEValPhePtPlrThTyGluysIleArgLumEt 317  

::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  

135453 CGCCGTGGCGCTCACCTTCTGCCTTCCTTGACAGATCCGCAAGATG 135404  

  

318 SerGIyvalSerProPhe 323  

||::|::|::|:  

135403 ATCGAGCCTCCGCCGCTAC 135386  

  

seq_name= gb_inl:AE003506  

  

seq_documentation_block:  

LOCUS AE003506 300994 bp DNA INV 04-OCT-2000  

DEFINITION Drosophila melanogaster genomic scaffold 142000013386053 section 23  

of 30 complete sequence.  

ACCESSION AE003506 AE002593  

VERSION AE003506.1 GI:7293355  

KEYWORDS HTG.  

SOURCE fruit fly,  

ORGANISM Drosophila melanogaster  

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  

Muscomorpha; Ephyridioidea; Drosophilidae; Drosophila .  

1 (bases 1 to 300994)  

Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
```

TITLE	The genome sequence of <i>Drosophila melanogaster</i>
JOURNAL	Science 287 (5461), 2185-2195 (2000)
MEDLINE	20196006
REFERENCE	2 (bases 1 to 300994)
AUTHORS	Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J
TITLE	Direct Submission
JOURNAL	Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
FEATURES	location/Qualifiers
source	1..300994
	/organism="Drosophila melanogaster"

MRNA  
 /db\_xref="taxon:7227"  
 /chromosome="x"  
 complement(join(<16198..16317,16386..16574,16635..16845,16915..17053,17129..17411,17526..17607,17676..17975,18107..>18483))  
 /gene="CG8527"  
 /product="CT24901"  
 /db\_xref="FLYBASE:FBan0008527"  
 /db\_xref="FLYBASE:FBgn0030844"  
 complement(<16198..16317,16386..16574,16635..16845,16915..17053,17129..17411,17526..17607,17676..17975,18107..>18483))  
 /gene="CG8527"  
 /note="CG8527 gene product"  
 /codon\_start=1  
 /db\_xref="FLYBASE:FBan0008527"  
 /db\_xref="FLYBASE:FBgn0030844"  
 /protein\_id="AAF4873.1"  
 /db\_xref="GI:729356"  
 /translation="MPQEKRRHPESRQRFLETVIFRSLIYOTKEFQNSLHGRVIAESGRIGEFPMKECTSGATYALYIINSLMEKQTNPTITGLDTPHNVPRITTVCEPAEAFHDKTYEKVYNTLANVDAQOMTPPLRLITSLNFENVDAKVLQSSIPONLDHILIREMAFEGCHIDCKNVFSCSKRDIIPCCHFEPIYHEHFCYAFNSRFKSTPELVKICGAPHDLETDKWALFEIPISTSRIFIFSNEEYFGSDFAQIDMSPRDLVEIRIKKMTYVTDARQSLIGORCIFEDEKLVNVPAYVFSCKMQRNNKAIKLCNPNPYPIPVDPWCSIKDPCDEFSNTNKKDCLQCLSCSKVPINDKIKMSDPESLGVLFELTPITRKREYVFGVDLVFSGTASLFLGFSLSGEYIYFTLLACQWKKRQELIEEKIRQEPNIDILSLKSNPRIQDTPSAVLKVPAAEESGPNPLGCGKRKEKQDSLMTKKPKVVPDGYGTSKTKNSWTYDHSQVAP"  
 complement(join(21273..21792,21868..21999,22058..22191,22257..22823,22889..23887,24168..24320,24398..25116,25191..25449))  
 /gene="CG8465"  
 /product="CT24743"  
 /db\_xref="FLYBASE:FBan0008465"  
 /db\_xref="FLYBASE:FBgn0030845"  
 complement(join(21273..21792,21868..21999,22058..22191,22257..22823,22889..23887,24168..24320,24398..25116,25191..25370,27861..28246))  
 /gene="CG8465"  
 /product="CT24757"  
 /db\_xref="FLYBASE:FBan0008465"  
 /db\_xref="FLYBASE:FBgn0030845"  
 complement(join(21273..21792,21868..21999,22058..22191,22257..22823,22889..23887,24168..24320,24398..25116,25191..25370,27861..28115,28453..28555))  
 /gene="CG8465"  
 /product="CT24733"  
 /db\_xref="FLYBASE:FBan0008465"  
 /db\_xref="FLYBASE:FBgn0030845"  
 complement(<21273..>28555)  
 /gene="CG8465"  
 /map="16B4-16B7"  
 /db\_xref="FLYBASE:FBan0008465"  
 /db\_xref="FLYBASE:FBgn0030845"  
 complement(join(21399..21792,21868..21999,22058..22191,22257..22823,22889..23887,24168..24320,24398..25116,25191..25370,27861..28092))  
 /gene="CG8465"  
 /note="CG8465 gene product [alt 1]"  
 /codon\_start=1  
 /db\_xref="FLYBASE:FBan0008465"  
 /db\_xref="FLYBASE:FBgn0030845"  
 /protein\_id="AAF4873.1"  
 /db\_xref="GI:729357"  
 /translation="MSTYFQVYIPTSKAGCEGVSQICGIAAVNIKPSNPASGSASVASGSPSSAASVQTVGNADDSAAATKEDDPYPPDSPLMWLFTFKSALDILRLRYKKA



|||||.....  
296686 GGGCTTTGGGGCTGTACAAAGCTTCTGCGCTGGATACGAATGG 296735  
301 hPrOTrPSeRMeTAlPhETrPLeuThrTyGluLysIleArgGluMet 317  
:::|||||.....  
296736 GCGCGTGGCGCTGACCTTCTGCTGCTCGTCCGTAACAGATCCGCAAGATG 296785  
318 SerGIyAlSeRPrOpHe 323  
|||||.....  
296786 ATCGGAGCCTCCGGCTAC 296803  
seq\_name: gb\_htg4:AC012693  
seq\_documentation\_block:  
LOCUS AC012693 79977 bp DNA HTG 03-NOV-1999  
DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*, In ordered  
pieces.  
ACCESSION AC012693  
VERSION AC012693.1 GI:6223614  
KEYWORDS HTG; HTGS\_PHASE2.  
ORGANISM Drosophila melanogaster  
fruit fly.  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 79977)  
AUTHORS Adams, M. and Venter, J.C.  
TITLE Direct Submission  
JOURNAL Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive,  
Rockville, MD, USA  
COMMENT This sequence was identified as CDM:1020955 by the submitter.  
For further information on this sequence you may e-mail to  
fly@celera.com.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
FEATURES  
Source 1. 79977  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
BASE COUNT 22591 a 17930 c 17947 g 21509 t  
ORIGIN  
alignment\_scores:  
Quality: 751.50 Length: 323  
Ratio: 3.447 Gaps: 4  
Percent Similarity: 67.492 Percent Identity: 46.130  
Alignment\_block:  
US-09-397-342-1 x AC012693/rev ..  
Align seg 1/1 to reverse of: AC012693 from: 1 to: 79977  
68 TyrATgGlyMeTAlArGTThrAlaLeuGlyIleIleGluGluGlyPh 84  
|||||.....  
35953 TACCGCGGCAATGAGTGGCCACCGCTTCGGGATTCGGCGAGAGAGCGGC 35904  
84 eLeuLysLeuTrpGlnGlyAlaThrProAlaIleTyrArgHisValT 101  
|||||.....  
35903 CCTGAAGCTGTGGCAGGCGGTACCGCGGCGCTCTACCGACACGTCGTCT 35854  
101 yr..... 101  
||  
35853 ATAGGTGGGCAATCCCAAACTCCTCATTTAGCATATAGATCAACCGCA 35804  
101 ..... 101  
35803 CAGCCAATTCGCGCAGCAGCTGTAGCCATCAGCCATTAGCTAGCAAG 35754  
102 ..... SerGIyAla 105  
|||||.....

35753 TTCTAACCAACTCTCCCTGCTCTCCCATGTCCATCCGACGCGGTGCA 35704  
105 rGMeTAlThrTyGTrGlnHisLeuArgGluAlaValPheGlyLysSerGlu 121  
|||||.....  
35703 GGATGTGAGCTAGCAGCACTGATGGCCAG...GAGTTACACACAGACGC 35657  
122 AspGlnHisTyRProLeuTrpLysSerValIleGlyIleMetAlaGlu 138  
:::|||||.....  
35656 ACCCAGGCTGCGCGTTGGAAAGTCGCGCTGTCGCGCGTACGACGCGCG 35607  
138 yAlIleGlnPheLeuAlaAsnProThrAspLeuValLysValGlnM 155  
:::|||||.....  
35606 AGCGTTCGCCAGTGGCTTCCTCCGCTGACATCTCTCAAGGTGGCAA 35557  
155 eTGlMeGluGlyLysArgLysLeuGluGlyLysProLeuArgPheArg 171  
:::|||||.....  
35556 TCCAATGAGAGGCCGACGAGCTGTGATGGCGGACCCAGCGGTGTGAC 35507  
172 GlyValHisHisAlaPheAlaLysIleLeuAlaGluGlyIleArgGlu 188  
:::|||||.....  
35506 TCCTGCGCATGCTCTCCGCGAGATCGTCGACGCGTGGCAATTAAGG 35457  
188 yLeuTrpAlaGlyTrpValProAsnIleGlnArgAlaAlaLeuValAsnM 205  
:::|||||.....  
35456 CCTGTGAAGGCGACACATCCGAATGTCCAGGACGCGCTGTGTCAATC 35407  
205 eTGIyAspLeuThrThyTrpAspThrValLysHisTyRLeuValLeuAsn 221  
:::|||||.....  
35406 TGGGCGACCTTAACCAACACACACACACACACACATGATATGATATGCG 35357  
222 ThrProLeuGluAspAsnIleMetThrHisGlyLeuSerLeuGlySse 238  
:::|||||.....  
35356 CTGCAGATGCCGCGACTGTACACACAGTGCAGCTGCGCTCGTTGTGC 35307  
238 rGlyLeuValAlaSerIleLeuGlyThrProAlaAspAlaIleLysSera 255  
:::|||||.....  
35306 CGGATCTGTGCGAGCATCATGAGCGACGCGCATGTGTGTGAAGACGC 35257  
255 rGlyLeuAsnGlnProArgAspLysGlnGlyArg..... 266  
|||||.....  
35256 GCATCATTAACACGCGCCACGACGAGATGCGCGTGAAGTACCAATGCG 35207  
267 ..... Gly.. 267  
35206 CGCAATTACTATTGGAACGCTTAATCACATATGTATTCTCGACGCGGC 35157  
268 LeuLeuTyRTrpSerSerThrAspCysLeuIleGlnAlaValGlnGly 284  
|||||.....  
35156 CTGCTATTCGCGGATTCGCTGACTGCTGCGACAGACAGAGGTTTGAAGGA 35107  
284 uGlyPheMetSerLeuTyRTrpGlyPheLeuProSerTrpLeuArgMet 301  
|||||.....  
35106 GGGCTTTTGGCGCTGTACAAAGCTTCTGCTCCCTGCGGTGATAGAAATG 35057  
301 hPrOTrPSeRMeTAlPhETrPLeuThrTyGluLysIleArgGluMet 317  
:::|||||.....  
35056 GCGCGTGTCTCCTCACCCTTGTGCTGTCTTGAAGATCCGCAAGATG 35007  
318 SerGIyAlSeRPrOpHe 323  
|||||.....  
35006 ATCGGAGCCTCCGGCTAC 34989  
seq\_name: gb\_htg6:AC017377  
seq\_documentation\_block:  
LOCUS AC017377 140468 bp DNA HTG 09-DEC-1999  
DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*, In ordered  
pieces.  
ACCESSION AC017377  
VERSION AC017377.1 GI:6553609  
KEYWORDS HTG; HTGS\_PHASE2.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster







REFERENCE  
AUTHORS

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 260550)

Amadiadis, G., Scherf, S.E., Li, P.-W., Hoskins, R.A., Galle, R.F., George, R.A., Lewis, S.E., Richards, S., Ashburner, M., Henderson, S.N., Sutton, G.G., Wortman, J.R., Vandeil, M.D., Zhang, Q., Chen, L.X., Brandon, R.C., Rogers, Y.H., Bazzel, R.G., Champe, M., Pfeiffer, B.D., Man, K.H., Doyle, C., Baxter, E.G., Helt, G., Nelson, C.R., Gabor Miklos, G.L., Abrell, J.F., Agbayani, A., An, H.J., Andrews-Pfannkoch, C., Baldwin, D., Ballwey, R.M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E.M., Beeson, K.Y., Benos, P.V., Bereman, B.P., Bhattacharya, D., Bolshakov, S., Borkova, D., Botchan, M.R., Bouck, J., Brokstein, P., Brotler, P., Butts, K.C., Busam, D.A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J.M., Cawley, S., Dahlke, C., Davenport, L.B., Davies, P., de Pablos, B., Delcher, A., Deng, Z., Mays, A.D., Dew, I., Dietz, S.M., Dodson, K., Doup, L.E., Downes, M., Dugan-Rocha, S., Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C., Ferrier, S., Fleischmann, W., Foster, C., Gabriellian, A.E., Garg, N.S., Gelbart, W.M., Glasser, K., Glodek, A., Gong, F., Gorrell, J.H., Gu, Z., Guan, P., Harris, M., Harris, N.L., Harvey, D., Helman, T.J., Hernandez, R., Houck, J., Hostin, D., Houston, K.A., Howland, T.J., Wei, M.H., Ibegwam, C., Jaitai, M., Kalush, F., Karpe, G.H., Ke, Z., Kennison, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., Laso, P., Lei, Y., Levitsky, A.A., Li, Y., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T.C., McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobarry, C., Morris, J., Mosnig, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L., Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusskern, D.R., Pacieb, J.M., Palazzolo, M., Pittman, G.S., Pan, S., Pollard, J., Pui, V., Reese, M.G., Reinert, K., Remington, R., Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I., Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C., Stapleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A., Weinstock, G.M., Weissenbach, J., Williams, S.M., Woodage, T., Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F., Zaveri, F.N., Zhu, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.

The genome sequence of *Drosophila melanogaster*  
Science 287 (5461), 2185-2195 (2000)

20196006

2 (bases 1 to 260550)

Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.  
Direct Submission  
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,  
Rockville, MD, USA

ATURES  
source

```

/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2L"
/complement(<174. .>467)
/gene="CG13998"
/product="CT33554"
/db_xref="FLYBASE:FBann013998"
/db_xref="FLYBASE:FBgn0040949"
/evidence=not_experimental
/complement(<174. .>467)
/gene="CG13998"
/map="26A-26A9"
/db_xref="FLYBASE:FBann013998"
/db_xref="FLYBASE:FBgn0040949"
/evidence=not_experimental
/complement(174. .467)
/gene="CG13998"
/note="CG13998 gene product"
/codon_start=1
/db_xref="FLYBASE:FBann013998"
/db_xref="FLYBASE:FBgn0040949"
/evidence=not_experimental
/protein_id="AAE52307.1"

```

## miRNA

```

/gb_xref="GI:7297038"
/translation="MAKTTWFLVLAIVLTNPWPFPFRTGAEVTSMSQFLDRHN
EGGDHSWSLLPTPNYSMPNDQYRFRFRGGRMDTFRSGRQOYPEYARV"
631. .1290
/gene="Ym26Ab"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/product="Cr25978"
/db_xref="FLYBASE:FBan0009046"
/gb_xref="FLYBASE:FBgn0003960"
<631. >1290
gene

```

## CDS

```

/db_xref="FLYBASE:FBan000904.6"
/db_xref="FLYBASE:FBan00033980"
/protein_id="AAF52308.1"
/db_xref="GI:7297039"
/translation="MAFNGHLLIGLVALSAVSSETTLOLOFTQGILLPAPLAENIR
SAAAGTGGAAPASYSAPAAAPAAQASAPAPAPYSAPAAAPASAPAPASAPAA
ASAPAPASAPASIPSPCKKNYLFSCQPLQPCSAQASYSAGAAQYVPO
AVPFWREL"
complement(<1708..>2253)
rRNA

```

gene

map="26a8-26a9"  
/db\_xref="FLYBASE:FBan0013997"  
/db\_xref="FLYBASE:FBgn0031754"  
evidence=not experimental  
complement(1708..2253)  
/gene="CG13997"  
/note="CG13997 gene product"  
/codon\_start=1  
/db\_xref="FLYBASE:FBan0013997"  
/db\_xref="FLYBASE:FBgn0031754"  
evidence=not experimental  
protein\_id="AAF52309.1"  
/db\_xref="GI:7297040"  
translation="MKKYLGFALLAGASLGGQAGLEASTGCGSTGCLRLSPSSSS  
VPSDYLARRRQOIYQVEVLYENNFSGGFGFGFGFGFGFGFGFGFGFG  
RFGFGPGGFG  
KNVYSCEAVINPVCYGSSY"  
complement(<3496..>3921)

gene

CDS

```

/note="Vm26Aa gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0009048
/db_xref="FLYBASE:FBgn0003979
/protein_id="AAF52310.1"

```

```
/db_xref="GI:7297041"
/translation="MKSEVCIALVFAAALASPNNVATGSGSVTTQDELECY
TGQGFIDLRLRKAYGGSSGSIIPAPCPKNLIFSCQPLIAPVCSAPAPSYG
SAGVSSPVATYVAPNYGVPOHOOLYSATVPQYGYQY"
complement(<6800..>7591)
/gene="CG9050"
/product="CT25990"
/db_xref="FLYBASE:FBan0009050"
/db_xref="FLYBASE:FBgn0031755"
complement(<6800..>7591)
/gene="CG9050"
/map="26A8-26A9"
/db_xref="FLYBASE:FBan0009050"
/db_xref="FLYBASE:FBgn0031755"
complement(6800..7591)
/gene="CG9050"
/note="CG9050 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0009050"
/db_xref="FLYBASE:FBgn0031755"
/protein_id="AAE52311.1"
/db_xref="GI:7297042"
/translation="MKLFKKLPPLFPASPCCGGCGGNNVAPAPSPSYGDCY
GVLPNGXIKPDTYNGRGGDAGYVAPAPAPAPAPAPAPAPAPAPAPAP
AAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP
IVTQPIILSTPLAKSSTQVETDQKVTAPAPPPPPAPAPVYEAPOSCYAPAP
PAPVATPSCOTPLRLIDQPYRVAPBELFEENYRLALASONTL"
complement(<9181..>11160)
/gene="CG13992"
/product="CT33548"
/db_xref="FLYBASE:FBan0013992"
/db_xref="FLYBASE:FBgn0031756"
/evidence=not_experimental
complement(<9181..>11160)
/gene="CG13992"
/map="26A8-26A9"
/db_xref="FLYBASE:FBan0013992"
/db_xref="FLYBASE:FBgn0031756"
/evidence=not_experimental
complement(9181..11160)
/gene="CG13992"
/note="CG13992 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0013992"

alignment_scores:
Quality: 705.00 Length: 344
Ratio: 2.866 Gaps: 5
Percent Similarity: 71.512 Percent Identity: 41.570

alignment_block:
US-09-397-342-1 x AE003612 ..
Align seg 1/1 to: AE003612 from: 1 to: 260550

5 GIJUGLUGLUGLArg.....LeuLeuProLeuThrGlnArgTr 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13615 GAGGAGAGAGAAAGCCCAAGTTGAGTACTGTGACCAACAAGAGAC 13664
17 proArgLaseLysPheLeuLeuSerLysAlaAlaThrValAlaG 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13665 TCCGCGCGTTGACTCTACCTCAGGCATTCGCTCCGCTCAGTCCG 13714
34 IuLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGln 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13715 AGATTGTGGCTATCCCTTCGATATGTGCAAGACCCGAATGACATCCG 13764
51 GIJGluAlaAlaLeuAlaArgLeuGlnAspGlyAlaArgGluSerAlaPr 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13765 GCGAGATCGCG...AGCAGGCTGGGT.....CAAGAAGCCAA 13799
67 cTyrrArgGluMetValArgThAlaLeuGlnLylleIleGluGluGluGlyP 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
13800 GTACCGAGGCTGTCTGGCCACTGCCATGGGCAATCGTCAAGGAGGGTGC 13849
84 heLeuLysLeuThrPrpGlnGlyValThrProAlaIleTyrAlaGlnHisVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13850 TCTGTAAAGCTCTACGCTGCGATATCCGCCATGCTTTCGCCACATCGCTC 13899
101 TyrSerGlyGlyArgMetValThrTyrGlnHisLeuArgGluValAlaPh 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13900 TTCAGTGGCATCAAAATGCTGACCTACGACTATGATCGGTGACAGATGAT 13949
117 eGlyLysSerGluAspGlnHisTyrProLeu....TrpLysSerValI 132
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13950 CGTGCCCGCAGGAGCGGAGCGCCAGCTATCTCTCGTGGATCTGCA 13999
132 IeGlyGlyMetMetAlaGlyValIleGlyGlnPheLeuAlaAsnProThr 148
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14000 TCAGTGGCGCTTACCGCGGCACTGCCAGGCTACTACAGATCCCAACC 14049
149 AspLeuValLysValIleMetGlnMetGlnGlyLysArgLysLeuGluG 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14050 GAGCTGATTAAATGCAATGCAATGAGGAGGAGCGGCGCTTAAGGGG 14099
165 yLysProLeuAlaGlyPheArgGlyValHisHisAlaPheAlaLysIleLeu 182
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14100 CAGGCTCCCGCATTCACAAATGCTCTCAGGCGCTGACTTCATTAACA 14149
182 IaGlyGlyLysLeuArgGlyLeuTyrPheAlaGlyTyrValProAsnIleGln 198
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14150 GAACGGGTGTGTGTGGTGTCTCTGGAAGGGAACGTTCACAAACATAG 14199
199 ArgAlaAlaLeuValAlaMet.....GlyAspLeu 205
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14200 CCGTGGCACTGTGACATAGTGAAGCTTAATAATCGGAGATTGGATCAG 14249
206 .....GlyAspLeu 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14250 TCAGGTACAGATATGTTCAATATACCCCTTCTTAAGGGAGCTCA 14299
209 hThrTyrAspThrValLysHisTyrLeuValLeuAsnThrProLeuGlu 225
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14300 GGTGTATGACTTTGGCAAGCGCTTGATGATGGCGATGTCGATCGTGG 14349
226 AspAsnIleMetThrHisGlyLeuSerSerLeuCysSerGlyLeuValAl 242
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14350 GATTAACCGAAGAGTCACTTCGTAGCCGCCATGACCGCGCGGTAGCGGA 14399
242 aserIleLeuGlyThrProAlaAspValIleLysSerArgIleMetAsn 259
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14400 TGCATTCGTAGCTTGCAGGAGCGAGCTGTCAAGTCCCGGATCATGAATC 14449
259 InProArgAspLysGlnGlyArgGlyLeuLeuTyrLysSerSerThrAsp 275
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14450 AGCCAAACGATGAGCAGGAGCGCGCATTCATTAAGAGGCTCCCTGATC 14499
276 CysLeuIleGlnAlaValGlnGlyLysGlyLysPheMetSerLeuTyrLysG 292
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14500 TGCCTTAAGTAGGTTGTAGAGAGAGGAGGCTTTTGGCCATGTACAAAGG 14549
292 yPheLeuProSerThrPheLysArgMetThrProTyrSerMetValPheTrp 309
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14550 ATTCAATCCATATATGATGATGCTGTGAGCTGCTCCGTGCTGTGTGGA 14599
309 euthrTyrGlyLysIleArgGluMetSerGly 319
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14600 TGACCTTGAGCAGATTTAGACGTTTCTGTGGC 14631

seq_name: gb_in2:CELK07B1
seq_documentation_block:
LOCUS CELK07B1 18557 bp DNA INV 14-MAY-1997
DEFINITION Caenorhabditis elegans cosmid K07B1.
ACCESSION AF003384 GI:2088817
VERSION
```

KEYWORDS  
SOURCE  
ORGANISM

Caenorhabditis elegans strain=Bristol N2.

Caenorhabditis elegans

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
Rhabditidae; Rhabditidae; Pelodierinae; Caenorhabditis.REFERENCE  
AUTHORS

1 (bases 1 to 18557)  
Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kersey, J., Kirsten, J., Laister, N., Latreille, P., Linton, J., Lloyd, C., McMurray, A., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifkin, L., Roopra, A., Saunders, D., Showkeen, R., Smailson, N., Smith, A., Sonhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkerson-Sprock, J., and Wohldman, P.

2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans

Nature 368 (6466), 32-38 (1994)

94150718

2 (bases 1 to 18557)

Pauley, A. and Gatlung, S.

The sequence of C. elegans cosmid K07B1

Unpublished (1997)

3 (bases 1 to 18557)

Waterston, R.

Direct Submission

Submitted (08-MAY-1997)

Submitted by:

Genome Sequencing Center  
Department of Genetics, Washington University,  
St. Louis, MO 63110, USA, and  
Sanger Centre, Hinxton Hall  
Cambridge CB10 1RQ, England  
e-mail: twenematode.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone

## NEIGHBORING COSMID INFORMATION:

The 5' cosmid is F44A2, 200 bp overlap; 3' cosmid is E02C12, 200 bp overlap. Actual start of this cosmid is at base position 31984 of CELF4A2; actual end is at 13615 of CELE02C12

The tandem region from 8582-10184; Assembly confirmed by restriction digest with BAMHI of cosmid DNA gives 5.2kb, database length is 5146 bases.

## NOTES:

Coding sequences below are predicted from computer analysis, using the program GeneFinder (P. Green and L. Hillier, ms in preparation).

Location/Qualifiers

1. 18557

/organism="Caenorhabditis elegans"

/strain="Bristol N2"

/db\_xref="taxon:6239"

/chromosome="V"

/clone="K07B1"

gene

CDS

/gene="K07B1.4"  
join(36..880,927..1101)

/gene="K07B1.4"

/codon\_start=1

/evidence=not\_experimental

/protein\_id="AB54240.1"

/db\_xref="GI:2088822"

/translation="MPHLGVEAPINPIPLARLQTLGALHFFETLFPVLVTPF

YMLTYLMLPLFLYGLMTLYDNKSPKRGAYMSWFORQTHSMYANFPVKLHTSDM

PEBNYLLIGHPIGITSMAFIFNATNGIGLIDTLPRIFPHLCTLGQPTWRNEMWG

LHGMIDCSRESIKVLEHKKKAVLVAGQAEALDAPGCHILTLKRRGFKVIA

LGTGAOLVPCSGENDIPNOAENPGSTIRGOFITMKRVLGSPAPARGVFNTF

GLTFAKPLNTVYAGPISVTKVFNPTQEOIDLHQYYMDRLHLEFHEHKTXYDSPTT

OLVIN"

complement(1305..2813)

/gene="K07B1.5"

complement(join(1305..1503,1551..1994,2547..2711,

2764..2813))

/gene="K07B1.5"

/codon\_start=1

/evidence=not\_experimental

/protein\_id="AB54237.1"

/db\_xref="GI:2088819"

/translation="MCDEQTKLLDLANGQPPGSRVLTIEQGSFASRLMVLKAPFT

LMCLSLVTFEATYFEGMLPVMWARTVWRRLTWVEGKLYRMLQSLIAYWGTAGDY

YEGDDVTYTYRDERIVLMCMNOSTADVTGIMTYLONKGVASRKTLMDWFRNTPF

GITGNHGDYFTIOGKATPDKELIRLKKHLHDFWDRRWVILPEEGFYTKRVSS

OSTGRKNGFPPLLYTTLPRMGAVQALIEEVPTEDESPDERRSNKLILQDTVGAI

REKKYVKGK"

4559..6150

/gene="K07B1.3"

join(4559..4799,4847..4998,5045..5359,5778..5905,

5955..6150)

/gene="K07B1.3"

/note="similar to the mitochondrial carrier family"

/codon\_start=1

/evidence=not\_experimental

/protein\_id="AB54239.1"

/db\_xref="GI:2088821"

/translation="MSSAVTSTSDSTAPGNSQFFKIATKVFITQVLPRTSIITYPY

FLSCATLVAETVYPLDITKRLIARNRKFTKGIMVQVYDDIRRGANALMTGV

APAITRHVYTYTIRMGVAGQIRLITFNKESVSPGLMSGLAOPASPTD

LVYVQMGELRLKLOKPLRYGATDCFEFLMGIMMPCORALLNWAD

IATYDSVKGILDNFELKDNMLTHAVASACAGIAAIVSLPDIYVTRMDQIRIED

AKMKHKKNTVDLKYKGVDCYIKIINKEGFSYKGLPSTYIRMAPWSLTFWVSEET

RKWTGASSP"

6548..8629

/gene="K07B1.2"

join(6548..6656,6730..6870,6920..7365,7733..7909,

7957..8135,8238..8299,8343..8629)

/gene="K07B1.2"

/note="strong similarity to the E. coli VISC protein

(SP:P25535)"

/codon\_start=1

/evidence=not\_experimental

/protein\_id="AB54238.1"

/db\_xref="GI:2088820"

/translation="MKLPGGTTCARNASSYDVTYVGGWGNAMAGSLGANKSFOS

KSVLLIDAGRSBSLAFKPGAPFNKRVNVSPTSIDTFKGLGWQDINSHRRKVRRL

FVDSCTSEIEERGOEVAEIIENDLVSLYKLAEVKADVDTAKEDCSIF

NLEENATYIKLENGDVETISLVRINFLNYFLFLEQIDAGVNSVRRASNDITTF

NYNQHGIIVATINETANKNETRANOREFTLGPVALLPESDYSGLTWSTSPEDARK

OLPSDFVDELNSALFSONNOIPIVNOTIFALNRNRPRTETFGRAEGTTPPHVTD

OKSRSRSPFGCNHASYITTCALIGDAHRNHPLAGGVNLGMSDVOILDKVLGDA

VRGADIGISTLYLREYDSSAOKRHNLPVWSVDLNLRTDAPATVAAARAGLNFNS

LGPVKNLNYLSAHR"

8582..10184

/note="tandem repeat"

repeat\_region

gene

complement(13165..13915)

CDS

complement(join(13165..13202,13462..13915))

CDS

/gene="K07B1.6"

CDS

/codon\_start=1



```

5706 AAAATATTTTAAAGTTCGCGGAATCTATTGTAATTTTATATTA 5755
225 .....GluAspAsnIleMethrHisGlyLeuSerS 235
5756 AAACATAACATATCAATTTTACAGAAATTTGGCTACATCCGCTGCTA 5805
225 erLeuCySerGlyLeuValAlaSerIleLeuGlyThrProAlaAspAl 251
5806 GTGCTGTGCTGGACTTCTGCTGCTGCTATATCATCTTCATCAGATGTC 5855
252 IleYsSerArgIleMetAsnGlnProArg..... 261
5856 GTTAAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5905
261 ..... 261
5906 GTTGTGATCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5955
262 .....AspLysGlnGlyValArgIleLeuLeuYrLysSerSerThrAs 275
5956 ATGCATAAGAAAATACCATGTGATCTA...TACAAAGCGCTGCTGCA 6002
275 pCysLeuIleGlnAlaValGlnGlyGlyPheMetSerLeuYrLysG 292
6003 TTGCTACTATAAAATATATCAAAAGAGATTTTTCACCTTACAAAG 6052
292 lYpLeuProSerTrpLeuArgMetTrpTrpSerMetValPheTrp 308
6053 GATTCCTCCATCTTATATTCGAAAGCGCTGCTGCTTACCTTCGCG 6102
309 LeuThrYrGlyLysIleArgIleMetSerGlyValSerProPhe 323
6103 GTTCCTACGAAGAAATCCGTAATGACCGCGCTTCAAGTTTC 6147
seq_name: gb_p14:AF155811

seq_documentation_block:
LOCUS AF155811 969 bp mRNA PRI 04-NOV-2000
DEFINITION Homo sapiens mitochondrial uncoupling protein 5 short form mRNA,
complete cds; nuclear gene for mitochondrial product.
ACCESSION AF155811
VERSION AF155811.1 GI:11094338
KEYWORDS
SOURCE
ORGANISM Homo sapiens
human.
REFERENCE
AUTHORS Yu,X.X., Mao,W., Zhong,A., Schow,P., Brush,J., Sherwood,S.W.,
ADAMS,S.H. and Pan,G.
TITLE Expression of uncoupling protein homologues, UCP4 and UCP5:
tissue-specific modulation by nutrition and temperature, and
evidence for UCP5 isoforms
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 969)
AUTHORS Pan,G.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Endocrinology, Genentech, Inc, 1 DNA Way,
Mail Stop-37, South San Francisco, CA 94080, USA
FEATURES
source
1. 969
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xq23-q25"
1. 969
/function="reduces mitochondrial membrane potential in
mammalian cells"
/note="UCP5; alternatively spliced"
/codon_start=1
/product="mitochondrial uncoupling protein 5 short form"
/protein_id="AAG29584.1"

```

```

/Db_xref="GI:11094339"
/translaton="MGFPGIILIFRVKFAATAVIHOKSTVSHMSGINKKPEYV
GIASIVAEKRGEPVDLTKRPIOVGOSIDARFEIKRGMFALRICKKEGVALYS
GIAPALROASVYTIKIGIQSKRLRFEVLEDELILINICGVSVISSTIANPTD
VKIRMOQSGSLTQSGMISGFIIDYQEGEGLRMVLPFAQAAIIVGVELPVYDT
KHLILSGMADTILTHFVSSFTGLAGALASNPVAVVTRMNOHAIIVGHVLYKGT
VDSILKMKHKEGFEALYKGFPMWMLRIGPWNILFITYEOLKRLQI"
BASE COUNT 258 a 187 c 240 g 284 t
ORIGIN

alignment_scores:
Quality: 557.50 Length: 296
Ratio: 2.617 Gaps: 5
Percent Similarity: 71.959 Percent Identity: 39.865

alignment_block:
US-09-397-342-1 x AF155811 ..
Align seg 1/1 to: AF155811 from: 1 to: 969

23 pheLeuLeuSerGlyCysAlaAlaThrValAlaGluLeuAlaThrPhePr 39
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
121 TTTGATATGCGCGGCTTGCCTGATCGTGGCTGAGTTGGACCTTCCC 170
39 oleuAspleuThrLysThrArgLeuGlnMetGlnGlyAlaAlaAlaLeu 56
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
171 TGTGACCTTACCAACACGACTTCAGGTTCAGGCCAACCAATGATG 220
56 laArgLeuGlyAspGlyAlaArgGluSerAlaProTyrArgGlyMetVal 72
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
221 CCCGTTTC.....AAAGATTAATATATAGAGGATGTTTC 255
73 ArgThrAlaLeuGlyIleIleGluGluGlyPheLeuLysLeuTrpGl 89
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
256 CATGCGCTTTCGCATCTGTAAAGAGAGGTGATGCTCTCTATTC 305
89 nGlyValThrProAlaIleTyrArgHisValValValYrSerGlyLysArg 106
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
306 AGCAATGCTCTCGCTGCTGCTAGACAAAGCATCATATGCGACCATTA 355
106 etValThrYrGlnHisLeuArgGluValAlaPheGlyLysSerGluAsp 122
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
356 TTGGATTTTACCAAGCTTGAAGCGCTTATGCTGAAGCTTTAGAGAT 405
123 GluHisTyrProLeuTrpLysSerValIleGlyGlyMetMetAlaGly 139
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
406 GAA.....ACTCTTTAATTATGATGCTGTGGGTAGTGTCAAGAGT 449
139 IleGlyGlnPheLeuAlaAsnProThrAspLeuValLysValGlnMetG 156
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
450 GATATCTTCCACTATATGCAATGCCATGCTTAAAGATTCGAATTC 499
156 lMetGluGlyLysArgLysLeuGluGlyLysProLeuArgPheArgGly 172
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
500 AGGCTCAAGAGC.....TTGTTCAAGG 525
173 ..ValHisHisAlaPheAlaLysIleLeuAlaGluGlyIleArgGl 188
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
526 AGCATGATTTGAAGCTTATCATATATACCAACAAGAGCAGCAGCGG 575
188 yLeuThrPheAlaGlyTyrValProAsnIleGlnArgAlaAlaLeuValAsn 205
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
576 TGTGTGAGAGGGTGTGTTCCAACTGCTGCTGCTGCTGCTGTTGAG 625
205 etGlyAspleuThrYrAspThrValLysHisTyrLeuValLeuAsn 221
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
626 GAGTAGAGCTACCATGATGATATTTACAAACATTTATATATGCA 675
222 ThrProLeuGluAspAsnIleMetThrHisGlyLeuSerLeuCysSe 238
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
676 GGAATGATGGCGATACATATTTAATCACTTCGTTCCAGCTTATACATG 725

```









OM of: US-09-397-342-1 to: N.Geneseq\_0401.\* out\_format : pfs  
Date: Apr 28, 2001 7:37 PM  
About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=cg2n2.1/USPTO.spool/US09397342/runat.27042001.170001.22942/app\_query.fasta.1.384  
-DB=N.Geneseq\_0401 -QFWT=faststep -SUFFIX=p2n.rng -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FCGAPOP=6.000 -FCGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELCP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blonsum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORMEXT -MINLEN=0 -MAXLEN=200000000  
-USER=US09397342@CGN1\_1199 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPHY  
-WAIT -THREADS=1

## Search information block:

Query: US-09-397-342-1  
Query length: 323  
Database: N.Geneseq\_0401.\*  
Database sequences: 678276  
Database length: 291890651  
Search time (sec): 113.230000

## score list:

Sequence	Strd Orig	ZScore	EScore	len	Documentation
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:A377139 +		1690.00	3208.05	8.3e-171	998 I HU
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:A14084 +		1690.00	3208.63	8.8e-171	1039 I C
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:A14085 +		1674.00	3176.10	5.7e-169	1248 I H
/SID56/gcgdata/geneseq/geneseq/NA2001.DAT:227721 +		1186.50	2240.75	7.1e-117	1726 I H
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:A46432 +		668.00	1254.67	6.0e-62	1069 I A
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C39993 +		668.00	1254.65	6.0e-62	1071 I A
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C45636 +		638.00	1199.75	1.2e-58	1240 I A
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C40326 +		638.00	1199.73	1.2e-58	1242 I A
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:A37719 +		560.00	1050.33	1.4e-50	876 I NC
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C89945 +		557.50	1044.85	3.9e-50	936 I HU
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C27995 +		557.50	1044.39	3.2e-50	978 I HU
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C27992 +		557.50	1044.22	3.2e-50	993 I HU
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C27996 +		557.50	1044.05	3.2e-50	1009 I H
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C298032 +		552.50	1034.68	6.5e-50	1674 I H
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C27999 +		552.50	1034.00	1.1e-49	1022 I M
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C46816 +		552.00	1029.73	2.0e-47	1103 I A
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C49402 +		531.50	992.30	2.5e-47	1124 I H
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C43096 +		515.00	992.28	2.5e-47	1227 I A
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C49389 +		515.00	993.37	1.0e-45	957 I A
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C76281 +		478.50	952.48	4.1e-45	828 I HU
/SID56/gcgdata/geneseq/geneseq/NA1998.DAT:C771710 +		478.50	891.02	1.1e-41	1220 I H
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:CX9435 +		476.50	890.08	1.2e-41	1930 I UC
/SID56/gcgdata/geneseq/geneseq/NA1998.DAT:C44595 +		476.50	888.25	2.6e-41	1612 I H
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:C119969 +		475.50	888.07	1.6e-41	939 I HU
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:CX08507 +		475.50	887.17	1.8e-41	1102 I H
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C229245 +		475.50	885.69	2.1e-41	1175 I H
/SID56/gcgdata/geneseq/geneseq/NA1998.DAT:C45602 +		475.50	885.53	2.2e-41	1193 I C
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C46055 +		475.50	885.53	2.2e-41	1193 I C
/SID56/gcgdata/geneseq/geneseq/NA1998.DAT:C772590 +		475.50	885.19	2.3e-41	1231 I H
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:C784254 +		473.50	887.38	3.5e-41	2340 I H
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:C219968 +		473.50	884.35	2.6e-41	930 I HU
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C50624 +		473.50	884.35	2.6e-41	930 I HU
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:CX9436 +		473.50	882.52	3.2e-41	1105 I HU
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C229233 +		473.50	882.52	3.2e-41	1105 I HU
/SID56/gcgdata/geneseq/geneseq/NA1998.DAT:C771712 +		473.50	881.61	3.6e-41	1204 I M
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:C84307 +		473.50	878.21	5.6e-41	1658 I M
/SID56/gcgdata/geneseq/geneseq/NA1998.DAT:C09078 +		472.50	879.25	4.9e-41	1255 I H
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:C82381 +		472.50	876.70	6.8e-41	1596 I H

seq_name: /SID56/gcgdata/geneseq/geneseq/NA2000.DAT:A37139					
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:A99716 +	472.50	876.70	6.8e-41	1596	
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:A52254 +	472.50	876.70	6.8e-41	1596	
/SID56/gcgdata/geneseq/geneseq/NA2000.DAT:C229244 +	472.50	874.96	8.5e-41	1882	
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:CX9436 +	472.50	874.92	8.5e-41	1888	
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:C207061 +	472.50	874.59	8.9e-41	1949	
seq_name: /SID56/gcgdata/geneseq/geneseq/NA2000.DAT:A37139					
seq_documentation_block:					
ID A37139 standard; cDNA: 998 Bp.					
XX A37139;					
XX AC A37139;					
XX AC A37139;					
DT 08-AUG-2000 (first entry)					
XX 08-AUG-2000 (first entry)					
DE Human UC94 cDNA sequence SEQ ID NO:405.					
XX Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;					
KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening;					
KW ss.					
XX Homo sapiens.					
OS Homo sapiens.					
PN WO200012708-A2.					
PD 09-MAR-2000.					
XX 09-MAR-2000.					
XX 01-SEP-1999; 99W0-US20111.					
XX 01-SEP-1998; 98US-0098716.					
PR 01-SEP-1998; 98US-0098749.					
PR 01-SEP-1998; 98US-0098750.					
PR 02-SEP-1998; 98US-0098803.					
PR 02-SEP-1998; 98US-0098821.					
PR 02-SEP-1998; 98US-0098843.					
PR 09-SEP-1998; 98US-0099536.					
PR 09-SEP-1998; 98US-0099596.					
PR 09-SEP-1998; 98US-0099598.					
PR 09-SEP-1998; 98US-0099602.					
PR 09-SEP-1998; 98US-0099642.					
PR 10-SEP-1998; 98US-0099741.					
PR 10-SEP-1998; 98US-0099754.					
PR 10-SEP-1998; 98US-0099763.					
PR 10-SEP-1998; 98US-0099792.					
PR 10-SEP-1998; 98US-0099808.					
PR 10-SEP-1998; 98US-0099812.					
PR 10-SEP-1998; 98US-0099815.					
PR 15-SEP-1998; 98US-0099816.					
PR 15-SEP-1998; 98US-0100385.					
PR 15-SEP-1998; 98US-0100388.					
PR 15-SEP-1998; 98US-0100380.					
PR 16-SEP-1998; 98US-0100384.					
PR 16-SEP-1998; 98US-0100627.					
PR 16-SEP-1998; 98US-0100651.					
PR 16-SEP-1998; 98US-0100662.					
PR 16-SEP-1998; 98US-0100664.					
PR 17-SEP-1998; 98US-0100683.					
PR 17-SEP-1998; 98US-0100684.					
PR 17-SEP-1998; 98US-0100710.					
PR 17-SEP-1998; 98US-0100711.					
PR 17-SEP-1998; 98US-0100919.					
PR 17-SEP-1998; 98US-0101071.					
PR 17-SEP-1998; 98US-0100930.					
PR 18-SEP-1998; 98US-0100848.					
PR 18-SEP-1998; 98US-0100849.					
PR 18-SEP-1998; 98US-0101014.					
PR 18-SEP-1998; 98US-0101068.					
PR 18-SEP-1998; 98US-0101071.					
PR 22-SEP-1998; 98US-0101279.					
PR 23-SEP-1998; 98US-0101471.					
PR 23-SEP-1998; 98US-0101472.					
PR 23-SEP-1998; 98US-0101474.					
PR 23-SEP-1998; 98US-0101475.					
PR 23-SEP-1998; 98US-0101476.					



```

134 1ymetMetalaGlyValIleGlyInPheLeuAlaAsnProThrAspLeu 150
    |||||||
427 GGATGATGCGCTGGTGTATTGGCCAGTTTTCACCAATCCACTGACCTA 476
151 VallysValGlnMetGlnMetGluGlyLysArgLysLeuGluGlyLysPr 167
    |||||||
477 GTGAAGGTTCGATGCAATGGAAGAAAAAGAACTGGAAGAAAAAC 526
167 OLeuArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAlaGlu 184
    |||||||
527 ATTCGATTCGTCGTGATCATGATTCGCAAAAATCTTACCTGAG 576
184 LysGlyIleArgGlyLeuThrPheAlaGlyThrPheAlaProAsnIleGluAla 200
    |||||||
577 GAGCAATACGAGGGCTTTGGGAGCTGGTACCAATATACAAAGACA 626
201 AlaLeuValAsnMetGlyAspLeuThrThrTyAspThrValLysHisTyr 217
    |||||||
627 GCACCTGGATATGGAGATTAAACCACTATGATACAGTGAACACTA 676
217 rLeuValLeuAsnThrProLeuGluAspAsnIleMetThrHisGlyLeu 234
    |||||||
677 CTGTGATTCGATACACCACTTGAGGACAAATATCATGACTCAGCGTTTAT 726
234 eTserLeuCysSerGlyLeuValAlaSerIleLeuGlyThrProAlaAsp 250
    |||||||
727 CAAGTTATGTTCTGAGCTGTAGCTCTATTCTGGGAGACACCGCAT 776
251 ValIleLysSerArgIleMetAsnGlnProArgAspLysGlnGlyArgG1 267
    |||||||
777 GTCAACAAAGCAGAAATATGATCAACACGAGTAAACAGGAAGGG 826
267 yLeuLeuTyIysSerSerThrAspCysLeuIleGlnAlaValGlnGly 284
    |||||||
827 ACTTTGATTAATCATCGACTGCTGCTGATTCAGGCTTTCAAGGTG 876
284 LuGlyPheMetSerLeuTyIysGlyPheLeuProSerThrLeuArgMet 300
    |||||||
877 AAGANTCATGAGTCATATATAAGGCTTTTACCATCTGGCTGAGATG 926
301 ThrProTyrSerMetValPheThrPheuThrTyGlyLysIleArgGlu 317
    |||||||
927 ACCCTTGCTCAATGCTGCTGCTGCTTACTTATGAAAAAATCAGAGAT 976
317 tSerGlyValSerProPhe 323
    |||||||
977 GAGTGGAGTCAGTCCATT 995
seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:AL14084
seq_documentation_block:
ID AL14084 standard; cDNA; 1039 BP.
AC AL14084;
XX
XX 21-JUL-2000 (first entry)
DE cDNA encoding human uncoupling protein UCP4.
XX
XX uncoupling protein; UCP4; expressed sequence tag; EST; human;
KW mitochondrial membrane; proton leakage; heat production; metabolic rate;
KM drug screening; obesity; stroke; trauma; sepsis; infection; ss.
XX
XX Homo sapiens.
XX
XX Key location/Qualifiers
XX CDS 40..1011
XX FT /*tag= a
XX FT /product= "Human UCP4"
XX
XX PN WO200017353-A1.

```

```

XX 30-MAR-2000.
PD
XX
XX 15-SEP-1999; 99WC-0521194.
PF
XX
XX 22-SEP-1998; 98US-0101279.
PR
XX 30-DEC-1998; 98US-0114223.
PR 16-APR-1999; 99US-0129674.
XX
XX (GETH ) GENENTECH INC.
XX
XX Adams S, Pan J, Zhong A;
PI
XX
XX WPI: 2000-292842/25.
DR
XX P-PSDB; Y81497.
PT
XX
XX New nucleic acid encoding human uncoupled protein-4, useful e.g for
PS identifying metabolic regulators for treatment of obesity.
XX
XX Claim 3; Fig 2; 80pp; English.
XX
XX This sequence represents cDNA encoding human uncoupling protein UCP4. The
CC human UCP4 cDNA (ATCC 203134) was isolated from a brain cDNA library
CC using a probe generated using primers AL14086-AL14087. These primers were
CC based on a UCP4 "from DNA" sequence (AL14085) derived from a number of
CC ESTs (expressed sequence tags) which were selected on the basis of
CC homology with human UCP3. The human UCP4 gene has been mapped to
CC chromosome 6p11.2-q12. UCP4 catalyses the leakage of protons through
CC the mitochondrial membrane, thus bypassing ATP synthase and thereby
CC reducing the efficiency of ATP synthesis. Modulation of UCP4 activity or
CC expression can therefore alter the metabolic rate and heat production
CC via modulation of ATP synthetic efficiency. UCP4 nucleic acids may be
CC used for recombinant production of UCP4 and as a source of primers and
CC hybridisation probes which may be used for the analysis of UCP4
CC expression, for screening for homologous sequences, and for chromosome
CC or gene mapping. They can be also be used to produce transgenic or
CC knockout animals for the development and screening of therapeutic agents,
CC as a source of antisense nucleotides, and in gene therapy for metabolic
CC disorders. The UCP4 protein can be used to screen for specific modulators
CC of activity and to raise antibodies which may be used therapeutically as
CC UCP4 antagonists or UCP4 targeting proteins, for affinity purification
CC of UCP4, and as immunoassay reagents for detecting UCP4 expression,
CC e.g., for the diagnosis of impaired neural activity or neural
CC degeneration. Agents that modulate UCP4 activity are used to control the
CC metabolic rate in mammals. UCP4 upregulators can be used to increase the
CC metabolic rate in order to treat obesity and the symptoms associated
CC with stroke, trauma, sepsis and infection.
XX
XX Sequence 1039 BP; 294 A; 215 C; 271 G; 259 T; 0 other;
XX

```

```

alignment_scores:
Quality: 1690.00 Length: 323
Ratio: 5.232 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment\_block:

US-09-397-342-1 x AL14084 ..

```

Align seg 1/1 to: AL14084 from: 1 to: 1039
1 MetSerValProGluGluGluArgLeuLeuProLeuThrGlnArgTr 17
  |||||||
40 ATGTCCGTCGCCGAGGAGGAGGAGGCTTTTGCCTGACCCAGATG 89
17 ProArgAlaSerLysPheLeuLeuSerGlyCysAlaAlaThrValAlaG 34
  |||||||
90 GCCCGGACGAGCAAAATTCATGTCGCGCTGCGCGCTACCGGCGCG 139
34 LuLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGln 50
  |||||||
140 AGCTAGCAACCTTTCCCTGATCTCACAACAAACGACTCCAAATGCA 189

```

```

51 GlyGluAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgLysSerAlaPr 67
190 GGAGAACCACTCTTCTGCTGGTGGAGAGCGTGCAGAGAAATCTGCCCC 239
67 CTATAGGAGGAAATGCTGGCAGACACCCCTAGGATCATTTGAAGAGAGAGCT 289
240 CTATAGGAGGAAATGCTGGCAGACACCCCTAGGATCATTTGAAGAGAGAGCT 289
84 heLeuLysLeuTrpGlnGlyValThrProAlaIleTyrArgHisValVal 100
290 TTCTAAAGCTTTGGCAAGGAGTACACCCCATTTTACAGACACAGTAGTG 339
101 TyrSerGlyLysArgMetValThrTyrGlnHisLeuArgLysValAla 117
340 TATTCGTGGAGGTGGAATGTCACATATGAAATCTCCGAGAGGTGTGTT 389
117 eGlyLysSerGlyAspGlyHisTyrProLeuTyrPlySerValIleGly 134
390 TGGCAAAAGTGAAGATGACATATCCCTTTGGAATCATGCTATTGGAG 439
134 LysMetAlaGlyValIleGlyGlnPheLeuAlaAsnProThrAspLeu 150
440 GGATGATGCTGCTGTTATGGCCAGTTTTCAGCAATCCAGTACGACCTA 489
151 ValLysValGlnMetGlnMetGlnGlyLysArgLysLeuGlnGlyLysPr 167
490 GTGAGAGTTCAGATGCAGAAATGGAAGGAAAGAACTGAGAGGAAACC 539
167 oLeuArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAlaGln 184
540 ATTGCATTTCTGCTGCTACATCATGATTTGCAAAAATTTAGCTGAG 589
184 LysIleArgGlyLeuTyrPalaGlyTyrValProAsnIleGlnArgAla 200
590 GAGGAATACAGAGGCTTTGGCAGCGCTGGTACCACATATACAAAGAGA 639
201 AlaLeuValAsnMetGlyAspLeuThrTyrAspThrValLysHisTyr 217
640 GCACCTGATGATATGGAGATTTAACCACTTATGATACACTGAAACCTA 689
217 rLeuValIleAsnThrProLeuGlnAspAsnIleMetThrHisGlyLeu 234
690 CTTGGATTCATACACCACTTACAGCAATATCATGACACGAGGTTTAT 739
234 eTrSerLeuGlySerGlyLeuValAlaSerIleLeuGlyThrProAla 250
740 CAAGTTATGCTCTGGACTGCTGATCTTCTGCTGGAAACACGACCGAT 789
251 ValIleLysSerArgIleMetAsnGlnProArgAspLysGlnGlyArg 267
790 GTGATCAAAAGCAGAAATTAAGATCAACCAAGATTAACAAAGAGAGG 839
267 yLeuLeuTyrLysSerSerThrAspCysLeuIleGlnAlaValGlnGly 284
840 ACTTTTGTAATATCATCCACTACCTGCTGATTCAGGCTGTCAAGGTG 889
284 LysIlePheMetSerLeuTyrLysGlyPheLeuProSerTrpLeuArgMet 300
890 AAGGATTCATGAGTCTATATAAGGCTTTTACCATCTTGGCGAGAGATG 939
301 ThrProTrpSerMetValPheTrpLeuThrTyrGlnLysIleArgLys 317
940 ACCCTTGGTCAATGCTGCTGCTTACTTATGAAGAAATCAGAGAGAT 989
317 tSerGlyValSerProPhe 323
990 GAGTGGAGTCAGTCCATTT 1008

```

seq\_name: /SID66/9cdata/geneseq/geneseqn/NA2000.DAT: A14085

seq\_documentation\_block:

ID A14085 standard; DNA: 1248 BP.  
XX

```

AC A14085;
XX
DT 21-JUL-2000 (first entry)
DE Human uncoupling protein UCP4 "from DNA" sequence, SEQ ID NO:5.
XX
KW Uncoupling protein; UCP4; expressed sequence tag; EST; UCP3 homologue;
KW human; chromosome 6p11.2-q12; ATP synthesis; energy efficiency;
KW mitochondrial membrane; proton leakage; heat production; metabolic rate;
KW drug screening; obesity; stroke; trauma; sepsis; infection; ds.
XX
OS Homo sapiens.
XX
PN WO200017353-A1.
XX
PD 30-MAR-2000.
XX
PE 15-SEP-1999; 99WO-US21194.
XX
PR 22-SEP-1998; 98US-0101279.
XX 30-DEC-1998; 98US-0114223.
PR 16-APR-1999; 99US-0129674.
XX
PA (GUTH ) GENENTECH INC.
XX
PI Adams S, Pan J, Zhong A;
XX
DR WPI: 2000-292842/25.
XX
PT New nucleic acid encoding human uncoupled protein-4, useful e.g for
XX identifying metabolic regulators for treatment of obesity.
XX
PS Example 1: Fig 7: 80pp; English.
XX
CC The invention relates to human uncoupling protein UCP4 (Y81497) and
CC cDNA encoding it (A14085). The human UCP4 cDNA (ATCC 203134) was isolated
CC from a brain cDNA library using a probe generated using primers
CC A14086-A14087. These primers were based on a UCP4 "from DNA" sequence
CC (A14085) derived from a number of ESTs (expressed sequence tags) which
CC were selected on the basis of homology with human UCP4. The human UCP4
CC gene has been mapped to chromosome 6p11.2-q12. UCP4 catalyses the leakage
CC of protons through the mitochondrial membrane, thus bypassing ATP
CC synthase and thereby reducing the efficiency of ATP synthesis. Modulation
CC of UCP4 activity or expression can therefore alter the metabolic rate and
CC heat production via modulation of ATP synthetic efficiency. UCP4 nucleic
CC acids may be used for recombinant production of UCP4 and as a source of
CC primers and hybridisation probes which may be used for the analysis of
CC UCP4 expression, for screening for homologous sequences, and for
CC chromosome or gene mapping. They can be also be used to produce
CC transgenic or knockout animals for the development and screening of
CC therapeutic agents, as a source of antisense nucleotides, and in gene
CC therapy for metabolic disorders. The UCP4 protein can be used to screen
CC for specific modulators of activity and to raise antibodies which may be
CC used therapeutically as UCP4 antagonists or UCP4 targeting proteins, for
CC affinity purification of UCP4, and as immunosay reagents for detecting
CC UCP4 expression, e.g., for the diagnosis of impaired neural activity or
CC neural degeneration. Agents that modulate UCP4 activity are used to
CC control the metabolic rate in mammals. UCP4 upregulators can be used to
CC increase the metabolic rate in order to treat obesity and the symptoms
CC associated with stroke, trauma, sepsis and infection. The present
CC sequence represents the human UCP4 "from DNA" sequence assembled from EST
CC sequences with homology to human UCP3.
XX
SQ Sequence 1248 BP: 357 A; 263 C; 301 G; 326 T; 1 other:

```

#### alignment\_scores:

Quality: 1674.00 Length: 323  
Ratio: 5.215 Gaps: 0  
Percent Similarity: 99.381 Percent Identity: 99.381

#### alignment\_block:

US-09-397-342-1 x A14085

Align seg 1/1 to: A14085 from: 1 to: 1248

```

1 MetSerValProGluGluGluArgLeuProLeuThrGlnArgTyr 17
|||||
26 ATGTCCTCCCGGAGGAGAGAGAGAGCTTGGCCGTGACCCAGAGATG 75
17 ProArgAlaSerLysPheLeuLeuSerGlyCysAlaAlaThrValAla 34
|||||
76 GCCCGGAGGAGCAATTCCTACTGCTCGGCTGCGGCTACCGCTGCGCG 125
34 IuLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGln 50
|||||
126 AGCTAGCAACCTTCCCTGGATCTCACAAAACCTGCATCCCAATGCAA 175
51 GlyGluAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgGlnSerAlaPr 67
|||||
176 GGAGAAACACCTCTTGGCTGCTGGAGAGCGGTGCAAGAGAACTGCCCC 225
67 OTYrArgGlyMetValArgThrAlaThrAlaGlyIleIleGluGluGluGlyP 84
|||||
226 CTATAGGGGAGATGTCGCCACAGCCCTAGGATCATTTGAGAGAGAGGCT 275
84 heLeuLysLeuTrpGlnGlyValThrProAlaIleTyrArgHisValVal 100
|||||
276 TTCTAAAGCTTTGGCAGAGAGTGACACCCGCCATTTCAGACAGCTAGTT 325
101 TyrSerGlyGlyArgMetValThrTyrGluHisLeuArgGluValValPh 117
|||||
326 ATTCTCTGAGGTGCAATGTCACATATGCAATCTCCGAGAGCTTGTCT 375
117 eGlyLysSerGluAspGlnHisTyrProLeuTrpLysSerValIleGlyG 134
|||||
376 TGGCAAAAGTGAAAGATGACATTATCCCTTGGAAATCAGTCATTGGAG 425
134 LymMetAlaGlyValIleGlyGlnPheLeuAlaAsnProThrAspLeu 150
|||||
426 GGATGAGGGCTGGTATTGGCCAGTTTTCAGCAATTCACAGCTAGCTTA 475
151 ValLysValGlnMetGlnMetGluGlyLysArgLysLeuGluGlyLysPr 167
|||||
476 GTGAGAGCTTCAGATGCAAAATGGAAAGAAAGAACTGGAGAGAAAC 525
167 OleuArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAlaGlu 184
|||||
526 ATTTCGATTTCGTGCTGCTACATCATGCAATTTGCAAAAATCTTACCTGAG 575
184 LysGlyIleArgGlyLeuTrpAlaGlyTyrValProAsnIleGlnArgAla 200
|||||
576 GAGCAATACGAAAGCTTTGGGCAAGCTGGGTACCCCAATATACAAAGACA 625
201 AlaLeuValAsnMetGlyAspLeuThrThrTyrAspThrValLysHisTyr 217
|||||
626 GCATCGGTGATATGAGAGATTTAACCACTTATATACAGGAAACACTA 675
217 rLeuValLeuAsnThrProLeuGlnAspAsnIleMetThrHisGlyLeuS 234
|||||
676 CTTCGTATGATACACCACTTGAGGACAAATATATGATGATCAGCGTTAT 725
234 erSerLeuCysSerGlyLeuValAlaSerIleLeuGlyThrProAlaAsp 250
|||||
726 CAAGTTATGTTCTGGAAGTGTACTTCTATTTGGGAGAACCCAGCGCAT 775
251 ValLysSerArgIleMetAsnGlnProArgAspLysGlnGlyArgGly 267
|||||
776 GTCATCAAAAGCAGAAATATGAAATCAACACAGATTAACAAGAGAGGG 825
267 YLeuLeuTyrLysSerSerThrAspCysLeuIleGlnAlaValGlnGlyG 284
|||||
826 ACTTTGTATTAATCATCGACTGCTGATTCAGGCTGTTCAGAGGTG 875
284 IuGlyPheMetSerLeuTyrLysGlyPheLeuProSerTrpLeuArgMet 300
|||||

```

```

876 AAGCATTCATGAGTCTATATATAAGCTTTTACCATCTTGCGTGAATG 925
301 ThrProTrpSerMetValPheTrpLeuThrTyrGluLysIleArgGlu 317
|||||
926 ACCCTTGGTCAATGCTTCTGCTTACTTATATGAAAAATCAGAGAGAT 975
317 tSerGlyValSerProPhe 323
|||||
976 GAGTGAAGTCAAGTTCATT 994

```

seq\_name: /SID6/gcgdata/geneseq/geneseqn/NA2001.DAT.F27721

seq\_documentation\_block:

ID F27721 standard; cDNA: 1726 BP.

XX F27721;

XX 28-MAR-2001 (first entry)

XX DE

XX DE

XX Human transport protein TPPT-21 coding sequence.

XX KW

KW Human; transport protein; TPPT; transport disorder; metabolic disorder;

KW neurological disorder; cardiovascular disorder; reproductive disorder;

KW immune disorder; cancer; ss.

XX OS

XX Homo sapiens.

XX PN

XX WO200078953-A2.

XX PD

XX 28-DEC-2000.

XX PF

XX 16-JUN-2000; 2000NC-US16668.

XX PR

XX 17-JUN-1999; 99US-0139923.

XX 10-AUG-1999; 99US-0148177.

XX 18-AUG-1999; 99US-0149357.

XX 28-OCT-1999; 99US-0162287.

XX PA

XX (INCY-) INCYTE GENOMICS INC.

XX PI

PI Lal P, Yang J, Yue H, Hillman JL, Tang YN, Bandman O, Burford N;

PI Baughn MR, Azimzal Y, Lu DAM, Au-Young J, Patterson C;

XX DR

XX P-PSDB; B60101.

XX DR

XX WPI: 2001-041424/05.

XX PT

PT Isolated polypeptide with a human transport protein sequence is useful

PT for the diagnosis, prevention and treatment of disorders associated

XX with the immune, reproductive and cardiovascular systems -

XX PS

PS Claim 5; Page 153-154; 165pp; English.

XX CC

CC The present invention provides the protein and coding sequences for 43

CC novel human transport proteins (designated TPPTs). These can be used in

CC the diagnosis and treatment of transport, metabolic, neurological,

CC reproductive, cardiovascular and immune disorders, and cell proliferative

XX disorders such as cancer.

XX SQ

SQ Sequence 1726 BP; 552 A; 302 C; 399 G; 473 T; 0 other;

alignment\_scores:

Quality: 1186.50

Ratio: 5.181

Percent Similarity: 93.852

Percent Identity: 92.623

alignment\_block:

US-09-397-342-1 x F27721 ..

Align seg 1/1 to: F27721 from: 1 to: 1726

95 IleyrArgHisVal..... 99

397 ATACATACCATATATACATATATACCAAGTTACTTTTATATATCT 446  
100 .....ValtyrSerglylyarGmetValhrrtyrGluHisLeuA 113  
TTTTTACCAGGTGTATTTGAGAGTGAATGGTCACATATGAACATCTCC 446  
113 rgluValValPheglylySergGluaspGluHisTyrrProleutPlys 129  
GAGAGCTGTGTGGCAAAAGAGAGATGAGCATTTATCCCTTTGGAAA 546  
130 SerValIleGlyGlymetLeuIleGlyValIleGlyGluPheLeuAAs 146  
547 TCAGTATTGGAGGATGATGGCTGTGTATTGGCCAGTTTATGCCCA 596  
146 nProthrAspLeuVallyValGlnmetGlnmetGluGlylySarGlySL 163  
597 TCCAATGACCTAGTGAAGTTTCAGATGCAATGAGAGAAAAGAAAC 646  
163 euGluGlylySproleuarGpheaTgGlyValhshIsaIaPheAlaLys 179  
647 TCGAAGGAAACCATTCGATTCGCTGCTGATCATCATCATTCGCAAAA 696  
180 IleLeuAlaGluGlylyIleLeuTrrPaIleGlyTrrPaIleProAs 196  
697 ATCTTAGCTGAAGAGGAATACAGAGGCTTTGGCAGGCTGGTACCACA 746  
196 nIleGlnArGAlaIleLeuValAsnmetGlyAspLeuThrThrTyraSP 213  
747 TATACAAGAGCAGCAGCTGTGAATATGGAGATTAAACCACCTATGATA 796  
213 hVallyshIstYrLeuValLeuAsnThrProleuGluAspAsnIleMet 229  
797 CAGTGAACACTACTGTGATTGAATACACACTTGAGGACATATCATG 846  
230 ThrHisGlyLeuSerSerLeuGlySerglyLeuValAlaSerIleLeuG 246  
847 ACTCAGGCTTATCAAGTTATGTTCTGGACGCTGGTCTTCTATTGGG 896  
246 yThrProIaaspValIleLysSerArgIleLeuAsnGlnProArgAsp 263  
897 AACACAGCCGATGTCATCAAAAGCAGATTAATGATCAACACGAGATA 946  
263 ysgGlnGlyArgGlyLeuLeuTyrrLysSerSerThrAspGlyLeuIleG 279  
947 AACAAAGAGGGGACTTTGTATTAATCATCAGCTGACTGATTGATCAG 996  
280 AlaValGlnGlyGluGlyPheMetSerLeuTyrrLysGlyPheLeuProse 296  
997 GCTGTTCAAGGAGGATTCATGACTATATAAAGGCTTTTATACCATC 1046  
296 rTrPLeuArgMetThrProTrpSerMetValPheTrPLeuThrTyrrGluL 313  
1047 TTGGCTGGAATGACCCCTTGTGCAATGGTGTCTGCTTACTATGAAA 1096  
313 ystIleArgGluMetSerglyValSerProPhe 323  
1097 AATCAGAGAGATGAGTGGAGTCAGTCATTT 1128  
seq\_name: /Sids6/gcdata/geneseq/geneseqn/NA2000.DAT:C46432  
seq\_documentation\_block:  
ID C46432 standard; DNA; 1069 BP.  
XX C46432;  
AC  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 50118.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KM protein identification; signal transduction pathway;  
KM metabolic pathway; promoter; termination sequence; ss.  
XX

OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139482.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.



PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## alignment\_scores:

Quality: 668.00 Length: 333  
Ratio: 2.760 Gaps: 9  
Percent Similarity: 72.673 Percent Identity: 42.342

## alignment\_block:

US-09-397-342-1 x C46432 ..

Align seg 1/1 to: C46432 from: 1 to: 1069

2 SerValProGluGluGluGluGluGluLeuLeuProLeuThrGluArgTyrPr 18  
||| ||| ..... ||||| ..... |||  
15 TCAAATCCGCTGCAAAATGAGCGG...AGCCGAGTACCAGAGAGACACC 61  
18 cArgAlaSerLysPheLeuLeuSerGlyCysAlaIaIaThrValAlaGluL 35  
| ..... ||||| ..... ||||| ..... |||||  
62 GACCGGAAACCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 111  
35 euAlaThrPheProLeuAspLeuThrLysThrArgLeuGluMetGlnGly 51  
: ||||| ..... ||||| ..... ||||| ..... |||||  
112 CAGTACGCTTCCGATTCATCTCACAAAGACTAGATGCGCTCCATGGT 161  
52 GUAAlaAlaLeuAla.....ArgLeuGlyAspGlyAlaArgGluSe 65  
: ..... : |||||  
162 TCGGATCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 194  
65 fAlaProTyrArgGlyMetValArgThrAlaLeuGlyIleIleGluGlu 81  
||| : ||||| ..... |||||  
195 .....GCTTTCGAGATCGTATGAGAGA 216



PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.

PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157785.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160747.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

alignment\_scores:                   Quality:   668.00                   Length:   333  
  Ratio:   2.760                   Gaps:    9  
Percent Similarity: 72.673           Percent Identity: 42.342

alignment\_block:  
US-09-397-342-1 x C39993   ..

Align seg 1/1 to: C39993 from: 1 to: 1071

2   SerValProGluGluGluGluArgLeuLeuProLeuThrGlnArgTyrPr 18  
   |||||                   |||||                   |||||  
17   TCAAATCCGGTCAAAATGAGCGG...AGCCGAGTGCACGAGACACCC 63  
18   cArgAlaSerLysPheLeuLeuSerGlyCysAlaIaIaThrValaIaGlu 35  
   |                   |||||                   |||||                   |||||  
64   GACCGAGACAGGATCGTCTGCTTCACTCTCTCGATGCTGCAGAGT 113  
35   euAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMeGlnGly 51  
   :::|||||                   |||||                   |||||                   |||||  
114   CAGTACGTTTCCGATGATCTCACAAAGACTAGATCAGATCCATGCT 163  
52   GuAlaIaLeuAla.....ArgLeuGlyAspGlyAlaArgGluSe 65  
   :::                   :::                   |||||  
164   TCGGATCCGCTTCTGCTGCGCATCGATGCTGC..... 196  
65   rAlaProTyrArgGlyMetValArgThrAlaLeuGlyIleIleGluGlu 81  
   |||||                   |||||                   |||||

```

197 .....GCTTCGAGCTGCTATCGAGA 218
82 .....GlutylPheLeuYsLeuTrpGlnGlyValThrProAla 94
219 TTGGAGAAAGAGAGAGATGTTGCTCTCAAGAGCTATCTCCGCGC 268
95 IleTyrArgHisValValTyrSerGlyGlyArgMetValThrTyrGluH1 111
269 ATATACAGACACCTGTTCTACACGCTATCAGAAATCATGTGATACGAGAA 318
111 sleuArgGluValValPheGlyYsSerGlu.....AspGluH1st 125
319 TTGTAAAGACTCATGCTC...AGATCTGAACACTACAAATGACGAGCTCTC 365
125 yrProLeu...TrpYsSerValIleGlyGlyMetMetAlaGlyValIle 140
366 TTCTCTCGGCACAAAGCTCTCTCTCGAGAGA...TTTCTGGGTATATA 412
141 GlyGlnPheLeuAlaAsnProThrAspLeuValYsValGlnMetGlnMe 157
413 GCTCAGCTAGTGGCTAGTCCAGCTGATTTGGTCAAGTGAATGCAAGC 462
157 tGluGlyValArgYsLeuGlyYsLeuYsProLeuArgPheArgGlyValH 174
463 AGATGTAAGTGTGTGAGCCAAAGCTGAAACCGAGGTACTCGGACCAA 512
174 IshaIaPheAlaYsIleLeuAlaGlnGlyGlyIleArgGlyLeuTrp 190
513 TCGAGGCTTTTACCAAAATCTCAATCAGAGAGATAAAGGTTATGG 562
191 AlaGlyTrpValProAsnIleGlnAlaAlaLeuValAsnMetGlyAs 207
563 AAGGCTTCTTCCAAATCAGACAGCATTTCTAGTAAATGAGAGA 612
207 PleuThrThrTyrAspThrValYsHisTyrLeuValLeuAsnThrPro 224
613 ACTAGCTGCTATGATCAGCCAAACACTTGTGATGATAGAGAAATGG 662
224 euGluAspAsnIleMetThrHisGlyLeuSerSerLeuCySerGlyLeu 240
663 CTGAGGATACATTTTGGCAGACCTTGTCTTATATATGCTCGGCTT 712
241 ValAlaSerIleLeuGlyThrProAlaAspValIleYsSerArgIleMe 257
713 GCCTCGCAAGTTGAGTGTCCAGCTGATGTGTGAAGACGAGATGAT 762
257 tsanGlnProArgAspYsGlnGlyArgGlyLeuLeuTyrYsSerSer 274
763 GAAC.....CAGGGTGAATAATGCTGTGTACAGAAATCTT 797
274 hrAspCysLeuIleGlnAlaValGlnGlyGlnGlyPheMetSerLeuTr 290
798 ACAGACTTGTGTGAGAGACGTTAAGTTGAGAGAAATGAGCTTGTGG 847
291 LysGlyPheLeuProSerTrpLeuArgMetThrProTrpSerMetValPh 307
848 AAAGGTTTCTTCCGACATGGGCAAGCTTGACCGTGCAGTTCGAGT 877
307 eTrpLeuThrTyrGlnYsIleArgGlnMetSerGlyValSerProPhe 323
898 TTGGGCTCTCTATGAGAGATTGACGCTGGCAGGAGATCTCTCCCTC 946
seq_name: /SID56/gcdata/geneseq/geneseqn/NA2000.DAT:C45636
seq_documentation_block:
ID C45636 standard; DNA: 1240 BP.
XX
XX C45636;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 47214.
DE
XX

```

```

KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134321.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 17-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.

```



```

65 rAlaProTyrArgGlyMetValArgThrAlaLeuGlyIleIleGluGlu 81
235 .....GCTTCGGAGTCGTATCGGAGA 256
82 .....GluGlyPheLeuIleuTyrPglngIValThrProAla 94
257 TTGGAGAAAGAGAGATGATGCTCTACAAAGGCTATCTCCGGCG 306
95 IleTyrArgHisValValTyrSerGlyArgMetValThrTyrGluH 111
307 ATTATCAGACACCTGTTCTACAGCCATATCATCTGATACGAGAA 356
111 sIeuArgGluValValPheGlyLysSerGlu.....AspGluHisT 125
357 TTGAAAGGATCATCTC...AGATCTGAACCTAACATACGAGCTCTC 403
125 yrProLeu...TrrLysSerValIleGlyLysMetMetAlaGlyValIle 140
404 TTCTCTCCGCCACAAAGCTCTCTCGGAGAGA...TTTCTCGGTATATA 450
141 GlyGln..Phe.....LeuAlaAsnProThrAspLeuValLysValGln 154
451 GCTCAGAAATGCAGATGATGCTAGTCACACTGATTTGTCAAGTGA 500
155 MetGlnMetGluLysArgLysLeuGluGlyLysProLeuArgPheAr 171
501 ATCCAAACCAATGATGATTGGTACGCCAAGCCCTGAACCGATGATCTC 550
171 gGlyValHisAlaPheAlaLysIleLeuAlaGluGlyGlyIleArg 188
551 GGGACCAATGAGAGCTTTTACCAAAATCTTACCAATCAGAAAGGTA 600
188 lyeuTrrAlaGlyTrrValProAsnIleGlnArgAlaAlaLeuValAsn 204
601 GGTATGGAAGGATGCTTCCAAACATCCAGAGACATTTCTAGTGAAT 650
205 MetGlyAspLeuThrThrTyrAspThrValLysHisTyrLeuValLeuAs 221
651 ATGGGGAACACTGCTATGATCATACGCCAAACACTTGTCTCATGA 700
221 nThrProLeuGluAspAsnIleMetThrHisGlyLysSerSerLeuLys 238
701 GAAGATTGCTGAGATACATTTTGGCCACACTCTTCTTATATATGT 750
238 erGlyLeuValAlaSerIleLeuGlyThrProAlaSpValIleLysSer 254
751 CGGGTCTTCTCGACAAAGTTGAGTTGTCACGCTGATGTGGAAGACG 800
255 ArgIleMetLsnGlnProArgAspLysGlnGlyArgGlyLeuLeuTyr 271
801 AGATGATGAAC.....CAGGTGAATAATCTGTGTACAG 835
271 sSerSerThrAspCysLeuIleGlnAlaValGlnGlyGlyPheMetS 288
836 AAATTTCTTACGACTTTTGGTGAAGACGTTAAGTTGAAGGATAAGAG 865
288 erLeuTyrLysGlyPheLeuProSerTrrLeuArgMetThrProTrrP 304
886 CTTTGTGGAAGCTTCTTCCGACATAGGCAAGGCTTGAACCGTGACG 935
305 MetValPheThrPheLeuThrTyrGluLysIleArgGluMetSerLys 321
936 TTCTGTCTTTGGGTCTCTATGAGAAGTTTACCTGCTGCGCAATCTC 965
321 rProPhe 323
986 TTCTCTC 992

```

```

AC C40226;
XX 17-OCT-2000 (first entry)
DT Arabidopsis thaliana DNA fragment SEQ ID NO: 27529.
XX
DE
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PN EP103405-A2.
XX
PD
XX
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 28-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0138119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.

```

```

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:C40226
seq_documentation_block:
ID C40226 standard; DNA: 1242 BP.
XX

```

```

PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151343.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
      Quality: 638.00      Length: 336
      Ratio: 2.626      Gaps: 11
Percent Similarity: 72.321      Percent Identity: 41.964

alignment_block:
US-09-397-342-1 x C40226 ..

Align seg 1/1 to: C40226 from: 1 to: 1242

2 SetValProGluGluGluArgLeuProLeuThrGlnArgTyrPr 18
||||| :||| :|||
57 TCAAAATCGGTGTAAGATGAGCGG...AGCCGAGTGCACAGACGAGCAGC 103
18 oryGalaSerLysPheLeuLeuSerGlyCysAlaAlaThrValAlaGlu 35
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
104 GACCGAGACCAAGATCTGCTTGGTTCACCTCTGCGATGGTTCCAGAGT 133
35 euAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGlnGly 51
::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

```

154 CAGTTACGTTTCCGATGTGATCTCACAAAGATGAGATGCAGTCCATGCT 203
52 GUAUAAALALEUAA.....ArgLeuGlyAspGlyAlaArgGluSe 65
204 TCGGGATCCGCTTCTGTGGCCGATCGGATGGC..... 236
65 TAAProTyrArgGlyMetValArgThrAlaLeuGlyIleIleGluGlu 81
237 .....GCTTCGGAGTCGTATCGAGA 258
82 .....GluGlyPheLeuLysLeuTrpGluGlyValThrProAla 94
259 TTGCGAAGAAAGAGAGATGATGCTCTCTACAAAGCTATCTCCGGCG 308
95 ILeTyrArgHisValAlaTyrSerGlyArgMetValThrTyrGluHis 111
309 ATTATCAGACACCTGTCTACAGCGCTATCAGATCATTTGGATTAGAGA 358
111 sLeuArgGluValValPheGlyLysSerGlu.....AspGluHisT 125
359 TTTGAAGAGACTCATCTC...AGATCTGAACATAACAATACGAGTCTC 405
125 yProLeu...TrpLysSerValIleGlyGlyMetMetAlaGlyValIle 140
406 TTCCCTCGCCACAAAGGCTCTCGTGGAGA...TTTCTGTGCTTTATA 452
141 GlyGln..Phe.....LeuAlaAsnProThrAspLeuValLysValGln 154
453 GGTCAAGAAATGCAGGTAGTGGCTACTCCAGCTGATTTGGTCAAACTGAGA 502
155 MetGlnMetGluGlyLysArgLysLeuGlyLysProLeuArgPheArg 171
503 ATGCAGAGAGATGATGATGGTGGAGCCAGGCTGAACCCAGGAGTACTC 552
171 gGlyValHisHisAlaPheAlaLysIleLeuAlaGluGlyIleArgG 188
553 GGGACCATTCGAGGCTTTTACCAAAATCTTACAAATCAGAGAGAGTAAAG 602
188 ILeuTrpAlaGlyTyrPValProAsnIleGlnArgAlaAlaLeuValAsn 204
603 GGTATGGAAGGTGCTCTTCCAAACATCCAGAGAGATTTCTATGAAAT 652
205 MetGlyAspLeuThrThrTyrAspThrValLysHisTyrLeuValLeuAs 221
653 ATGGAGAACTAGCTTGTGATGATCAGCGCAAAACATTGTATGATGAA 702
221 nThrProLeuGlnAspAsnIleMetThrHisGlyLeuSerSerLeuGly 238
703 GAAGATTCGCTGAGATTAACATTTTCCGACACACTCTGCTTATATAAT 752
238 erGlyLeuValAlaSerIleLeuGlyThrProAlaAspValIleLysSer 254
753 CCGGCTTCTGTCGACAAAGTTGAGTGTGCCAGTGTGTTGAAGAG 802
255 ArgIleMetAsnGlnProArgAspLysGlnGlyArgGlyLeuLeuTyrIly 271
803 AGGAGATGAGAAC.....CAGGCTGAATGCTGTATACG 837
271 sSerSerThrAspCysLeuIleGlnAlaValGlnGlyLysGlyPheMetS 288
838 AAATCTTACACACTTTTGGTGAACACGCTTAGATTGAAGAAATAAAG 887
288 erLeuTyrLysGlyPheLeuProSerTyrPLeuArgMetThrProTyrP 304
888 CTTTGTGGAAGAGTTTCTCCGACATGGGCAAGCTTGAGACGTCGAC 937
305 MetValPheTrpLeuThrTyrGluLysIleArgGluMetSerGlyValSe 321
938 TTCGCTTTGGTCTCTATGAGAAATTAGACTGTGTCGAGGAATCTC 987
321 rProPhe 323
988 TTCCCTTC 994

```

```

seq_name: /SID56/gcgdata/geneseq/geneseqn/NA2000.DAT.A37719
seq_documentation_block:
ID A37719 standard; cDNA; 876 BP.
XX
AC A37719;
XX
DT 22-NOV-2000 (first entry)
XX
DE Novel human protein coding sequence #1.
XX
KW Novel human protein; NHP; uncoupling protein; diagnosis; therapy;
KW metabolism; thermogenesis; obesity; cachexia; ss.
XX
OS Homo sapiens.
FH
FT Key 1..876
FT CDS /*lag- a
FT /product= NHP#1
XX
XX WO200047617-A1.
XX
XX 17-AUG-2000.
XX
XX 09-FEB-2000; 2000WO-US03425.
XX
XX 09-FEB-1999; 99US-0119228.
XX
XX 08-OCT-1999; 99US-0158458.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Turner CA, Mathur B, Zambrowicz B, Sands AT;
XX
XX WPI: 2000-558191/51.
XX
XX P-PSDB: Y90322.
XX
XX Nucleic acids encoding human uncoupling proteins useful for the study
XX and modulation of conditions such as obesity and cachexia -
XX
XX Claim 1; Page 25; 31pp: English.
XX
XX This sequence encodes a novel human protein (NHP) of the invention. The
XX NHP's of the invention are uncoupling proteins. The DNA and the encoded
XX protein may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate expression of the protein. The DNA
XX may be administered to treat diseases by rectifying mutations or
XX deletions in a patient's genome that affect the activity of NHP by
XX expressing inactive proteins or to supplement the patient's own production
XX of NHP polypeptides. Additionally, the DNA may be used to produce the
XX protein. Conversely, complementary sequences and antisense nucleic acid
XX molecules may be administered to down regulate protein expression by
XX binding with the cells own NHP genes and preventing their expression. The
XX DNA and the antisense molecules may also be used as DNA probes in
XX diagnostic assays to detect and quantitate the presence of similar
XX nucleic acid sequences in samples, and hence which patients may be in
XX need of restorative therapy. They may also be used to study the
XX expression and function of the NHP's and their role in metabolism by the
XX creation of transgenic and/or gene knock-out animals. The NHP
XX polypeptides may be used as antigens in the production of antibodies
XX against NHP and in assays to identify modulators (agonists and
XX antagonists) of NHP expression and activity. Anti-NHP antibodies and
XX NHP antagonists may also be used to down regulate NHP expression and
XX activity. Anti-NHP antibodies may also be used as diagnostic agents for
XX detecting the presence of NHP polypeptides in samples. Processes and
XX conditions associated with NHP polypeptides in samples. Processes and
XX thermogenesis, obesity and cachexia.
XX
XX Sequence 876 BP; 237 A; 184 C; 229 G; 226 T; 0 other;

```

alignment\_scores:



Quality: 560.00 Length: 295  
Ratio: 2.605 Gaps: 4  
Percent Similarity: 72.881 Percent Identity: 38.983

Alignment\_block:

US-09-397-342-1 x A37719 ..

Align seg 1/1 to: A37719 from: 1 to: 876

```

23 PheLeuLeuSerGlyCysAlaAlaThrValAlaGluLeuAlaThrPhePr 39
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 TTTGTGACGGGGGCTGCTCCATCATCTGCTGAGTGGCGATTTCC 74
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
39 OleuLeuPheThrThrArgLeuGluMetGluGluAlaAlaLeuA 56
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
75 AATGATTTAAACAGACAGCGCTCCAGATTCAGGCCAGC..... 117
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
56 1aArgLeuGlyAspGlyAlaArgGluSerAlaProTyrArgGlyMetVal 72
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
118 ..... AATGATGCAGAAATTTAGGAAATTTAGATCCGAGATGTG 159
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73 ArgThrAlaLeuGlyTleIleGluGluGlyPheLeuLysLeuTrpGl 69
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
160 CACGATAGTAGAGATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 209
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
89 nglValAlaProAlaIleTyrArgHisValValTyrSerGlyArgm 106
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
210 GGGGATTCGCCCCCGGATGTTACGCCAGGACATCCATGACCATCAAGA 259
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
106 etValThrTyrGluHisLeuArgGluValAlaPheGlyLysSerGluAsp 122
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
260 TAGGACTTACAGAGCTTGAAGCGCATTTATTCATGAGCCAGCAAGAT 309
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123 GluHisTyrProLeuTyrLysSerValIleGlyMetMetAlaGlyVa 139
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
310 GAAACTCTACCGATA..... AATGTATGTGTGGAATTTCTGTGAGAT 353
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
139 1IleGlyInPheLeuAlaAsnProThrAspLeuValLysValGlnMetG 156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
354 CATATCTTCAACCATTTGCTTAATCCAACTGATTTTGAATAATTCGAGATC 403
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
156 1MetGluGlyLysArgLysLeuGluGlyLysProLeuArgPheArgGly 172
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
404 AAGCCCA... AGCAACACATCAAGA:..... GGA 432
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
173 ValHisAlaPheAlaLysIleLeuAlaGluGlyIleArgGlyLe 189
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
433 ATGATAGGCACTTCATGACATTTACCAAGCAAGAGGAGACAGAGACT 482
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
189 uTPAlaIleTyrPValProAsnIleGlnArgAlaAlaLeuValAsnMetG 206
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
483 GTGGAAGGTTGTGCTTACTGCGCAGAGGCGCTGCTTGTGTGTGTTG 532
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
206 1AspLeuThrThrTyrAspThrValLysHisTyrLeuValLeuAsnThr 222
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
533 TGGACCTCCGCTATATGACATACCAAGACATCTTATTTCTTCAGGC 582
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
223 ProLeuGluAspAsnIleMetThrHisGlyLeuSerSerLysCysSerG 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
583 CTGATGGAGACACTGTGATACCACTTCCTCAGCTTCACCTGAGG 632
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
239 yLeuValAlaSerIleLeuGlyThrProAlaAspValIleLysSerArgI 256
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
633 TCTGTACAGGGGCGCTGAGCCCTCAACCCGTGTGATTTGTGAGACACGTA 682
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
286 1eMetAsnGluProArgAspLysGlnGlyArgGlyLeuLeuTyrLysSer 272
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
663 TGATTAATCAGAGAGTGTCTGAGATGCGAGATGTTCTGCTACACAGGA 732
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
273 SerThrAspCysLeuIleGlnAlaValGlnGlyGluGlyPheMetSerLe 289
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
733 ACCCTGATTTGCTTACAGACATGAGAAATGAAGGTTTGTGCTCT 782

```

```

289 uTyrLysGlyPheLeuProSerTyrPheArgMetThrProThrSerMetV 306
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
783 CTATTAAGGCTTTTGGCCAAATGCTTGACACTTGTCTTGAGATATCA 832
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
306 1aPheThrLeuThrTyrGluLysIleArgGluMet 317
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
833 TTTTCTTGTGACATACGACGAGCTTGAGAAATG 867

```

seq\_name: /SID56/gcgdata/geneseq/geneseqn/NA2000.DAT:289945

seq\_documentation\_block:

ID 289945 standard; cDNA; 936 BP.

AC 289945;

DT 05-MAY-2000 (first entry)

XX Human uncoupling protein 4 (UCP-4) nucleotide sequence.

XX Uncoupling protein 4; UCP-4; human; mitochondrial proton transport;

KW increase thermogenesis; obesity treatment; fat decrease; diabetes; fever;

KW respiratory ATP synthesis; malignant hyperthermia; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 37..909

FT /tag= a

FT /product= UCP\_4

FT /note= "Uncoupling protein 4"

XX NO200004037-A1.

XX 27-JAN-2000.

XX 13-JUL-1999; 99MO-US15861.

XX 14-JUL-1998; 98US-0092737.

XX (AMYL-) AMYLIN PHARM INC.

XX Albrandt K, Beaumont K, Young AA;

XX WPI; 2000-161095/14.

XX P-PSDB; Y78511.

XX Novel uncoupling protein 4 (UCP-4) used to treat obesity, diabetes,

XX fever, malignant hyperthermia and fever -

XX Claim 5; Fig 1; 72pp; English.

This sequence represents the human uncoupling protein 4 (UCP-4) nucleotide sequence. UCP-4 is expressed in various tissues including brain, heart, pancreas and muscle tissue. UCP-4 may have a role similar to other uncoupling proteins in mediating mitochondrial proton transport. The UCP-4 protein has antidiabetic, anorectic and antihypertensive properties. The UCP-4 protein can be used for increasing thermogenesis, treating obesity, decreasing the amount of fat, or treating diabetes. Antisense nucleic acids to UCP-4 are used for decreasing thermogenesis or respiratory ATP synthesis. The UCP-4 protein is used to screen for compounds which bind to or modulate the activity of UCP-4. Compounds that increase activity of UCP-4 are used for treating conditions or disorders that can be ameliorated by increasing the level of thermogenesis in a subject (e.g. obesity or diabetes). Compounds that decrease UCP-4 activity are used for treating conditions or disorders that can be ameliorated by decreasing the level of thermogenesis in a subject (e.g. malignant hyperthermia or fever). Modulatory compounds can also be used to regulate insulin secretion, and for increasing energy-maximizing responses or increasing energy-minimizing responses.

Sequence 936 BP; 252 A; 183 C; 231 G; 268 T; 2 other;



SQ Sequence 978 BP; 261 A; 189 C; 243 G; 285 T; 0 other;  
alignment\_scores:  
Quality: 557.50 Length: 296  
Ratio: 2.617 Gaps: 5  
Percent Similarity: 71.959 Percent Identity: 39.865  
alignment\_block:  
US-09-397-342-1 x A27995 ..  
Align seg 1/1 to: A27995 from: 1 to: 978  
23 PheLeuLeuSerGlyCysAlaAlaThrValAlaGluLeuAlaThrPhePr 39  
130 TTTGTATATGGCGCGCTGCTATCGTGCAGCTTTGGACTTCCC 179  
39 oLeuAspLeuThrLysThrArgLeuGlnMetGlnGlyAlaAlaLeuA 56  
180 TGTGGACCTTACCAAAACAGACTTCAGTTCAAGGCCAAGCATTCATG 239  
56 IaArgLeuGlyAspGlyAlaArgGluSerAlaProTyrArgGlyMetVal 72  
230 CCGGTTTC.....AAGAGATAAATATAGAGGAGATGTTTC 264  
73 ArgThrAlaLeuGlyLeileGluGluGlyPheLeuLysLeuTPgl 89  
265 CATGGCGTGTTCGATCTGTAAAGAGAGAGGTATGGCTCTATATTC 314  
89 nGlyValThrProAlaIleTyrArgHisValAlaTyrSerGlyGlyArgm 106  
315 AGAATTTGCTCGCGTGTCTAGACAGCATCATGCGACCATTAATAA 364  
106 eValAlnThrTyrGluHisLeuArgGluValAlaPheGlyLysSerGluasp 122  
365 TTGGATTATTCACCAAGCTTGACGCGTTATTCGTAGAACCTTGAAGAT 414  
123 GlnHisTyrProLeuTyrLysSerValIleGlyLysMetMetAlaGlyVa 139  
415 GAA.....ACTCTTTTAATTAATATGATCTGTGGGCTAGTCAGAGAT 458  
139 IleGlyGlnPheLeuAlaAsnProThrAspLeuValLysValGlnMetG 156  
459 GATATCTTCACATATAGCCAATCCACCGCATGTTCTAAAGATTGCAATGC 508  
156 lMetGluGlyLysArgGlyLysLeuGluGlyLysProLeuArgPheArgGly 172  
509 AGGCTCAAGAGAGC.....TTGTTCCAAGGCG 534  
173 ...ValHisHisAlaPheAlaLysIleLeuAlaGluGlyLysIleArgG 138  
535 AGCATGTTTGGAAAGCTTTATCGATATATACCAACAACAGAGGACAGGG 584  
188 yLeuTyrPalaGlyTyrValProAsnIleGlnArgAlaAlaLeuValAsnm 205  
585 TCTGTGAGAGGGGTGTGCTCACTGCTGCGTGCATGCGCATGTTGTAG 634  
205 eTcLysPleuThrThrTyrAspThrValLysHisTyrIleValLeuAsn 221  
635 GAGTAAAGCTACACGCTATGATATTAAGTAAAGCATTTAAATATGTCTCA 634  
222 ThrPheLeuGluAspAsnIleMetThrHisGlyLysSerSerLeuYsse 238  
685 GGAATGATGGCGCATACAAATTTTAACCTACTGCTTCCAGCTTTACATG 734  
238 rGlyLeuValAlaSerIleLeuGlyThrProAlaAspValIleLysSerA 255  
735 TGTGTGGCTGGGGCTGTGCTCAACCGGTTGATGTGCTTCAAGATC 784  
255 rGlyLeuTangInProArgAspLysGlnGlyArgGlyLeuLeuTyrLys 271  
785 GCATGATGAACACAG...AGGGCAATCGTGGACATGTGATCTCTATTAAG 831

272 SerSerThrAspCysLeuIleGlnAlaValAlaGlnGlyLysPheMetse 288  
832 GGCACGTGTATGATATTAAAGATGTGCAACATGAGGCGTTTTC 881  
288 rLeuTyrLysGlyPheLeuProSerThrLeuArgMetThrProTyrSerm 305  
882 ACTCTTAAAGGATTTGGCCAAACGCTGCTGACCTGACCTGGAACA 931  
305 eValPheThrPleuThrTyrGluLysIleArgGluMet 317  
932 TCATTTTATTTATACATACGACGCTAAAGAGGCTT 969  
seq\_name: /SID6/gcdata/geneseq/geneseqn/NA2000.DAT:A27992  
seq\_documentation\_block:  
ID A27992 standard; DNA; 993 BP.  
XX  
AC A27992;  
XX  
DT 29-AUG-2000 (first entry)  
XX  
DE Human uncoupling protein (UCP5) nucleotide sequence.  
XX  
XX  
KW Uncoupling protein 5; UCP5; metabolism; chromosome 10q23-25; H+ leak;  
KW metabolic rate; obesity; stroke; trauma; burn trauma; sepsis; infection;  
KW human; ds.  
XX  
OS Homo sapiens.  
XX  
FH  
FT CDS  
FT  
FT sig\_peptide  
FT  
FT mat\_peptide  
FT  
PN WO200032624-A2.  
PD 08-JUN-2000.  
XX  
PF 03-NOV-1999; 99WO-US25947.  
XX  
PR 30-NOV-1998; 98US-0110286.  
PR 16-APR-1999; 99US-0129583.  
PR 15-JUL-1999; 99US-0143886.  
XX  
PA (GENE ) GENENTECH INC.  
XX  
PI Adams S, Pan J;  
XX  
DR WPI: 2000-412284/35.  
XX P-PSDB: Y94665.  
PT Isolated nucleic acid encodes human uncoupling protein 5 useful in  
PT diagnostic assays and treatment of obesity, stroke, trauma, sepsis and  
PT infection -  
XX  
PS Claim 2: Fig 1; 90pp: English.  
XX  
CC This sequence represents a human uncoupling protein 5 (UCP5) encoding  
CC DNA sequence. UCP5 is involved in metabolism, and it may be involved in  
CC catalyzing H+ leak, and therefore be involved in energetic inefficiency  
CC in vivo. The present invention relates to human and murine UCP5  
CC nucleotide and protein sequences. There are three isoforms of human UCP5,  
CC hucP5L, hucP5S, and two isoforms of murine UCP5, mUCP5L and  
CC mUCP5S. The human UCP5 gene is located on chromosome 10q23-25. The  
CC nucleic acids encoding UCP5 can be used as hybridization probes, in  
CC chromosome and gene mapping, for the generation of antisense RNA and DNA  
CC and in the preparation of recombinant UCP5 proteins. UCP5 nucleic acids  
CC can be used in gene therapy for regulation of metabolic conditions.



Sequence 1009 BP; 267 A; 197 C; 249 G; 296 T; 0 other;

Quality:	557.50	Length:	296
Ratio:	2.617	Gaps:	5
Percent Similarity:	71.959	Percent Identity:	39.865

09-397-342-1 x A27996

Align seg 1/1 to: A27996 from: 1 to: 1009

23 PheLeuSerGlyCysAlaIleThrValAsnIleuAlaIsthPheR 39  
121 TTTGTATATGGCGGCTTTCGCCCTATACGGCTGAAGTTTGAGCATTTCC 170  
39 OluAspLeuThrLysThrArgLeuGlnMetGlnGluAlaAlaLeuA 56  
171 TGTTGACTTACCAAAACACGACTTCAGATTCAAGCCAAACATTGATG 220  
56 IaArgLeuGlyAspGlyAlaArgGluSerAlaProTyrrArgLymeAl 72  
221 CCCGTTTC.....AAGAAGATAAATATATGAGGAGATGTC 255  
73 ArgThrAlaLeuGlyIleIleGlnGluGlnGlyPheLeuLysLeuTrpL 89  
256 CATCCGCCTGTTCCGATCTGTAAAAAGAGAGGTGTATTTGGCTCATTC 305  
89 nGlyValThrProAlaIleTyrrArgHisValValYrrSerGlyLyrGm 106  
306 AGGAATTCCTCTCGCTGCTTAAGACAAGACATCATATGCCACCATTTAAA 355  
106 eValIThrTyrgLnuHISLeuArGluValAlaPheGlyLysSergLuasp 122  
356 TTGGGATTTACCAAAGCTTGAAGCGCTTATTCGTGAAGCTTAGAAGAT 405  
123 GlnHisTyrrProLeuTrpLysSerValIleGlyGlyMeMetAlaGlyA 139  
406 GAA.....ACCTGTTTAAATTAATGTATCTGTGGGTGAGTGCAGAGCT 449  
139 ILeGlyGlnPheLeuAlaAsnProThrAspLeuValLysValGlnMeG 156  
450 GATATCTTCCACATATAGCCAATCCACCGCATCTTCTAAAGATTCCATGC 499  
156 lmeGtGlnGlyLysArGLysLeuGlnGlyLysProLeuArGPheArgGly 172  
500 AGGCTCAAGGAAGC.....TTGTTCCAAGG 525  
173 ...ValHSHSLaPheAlaLysIleLeuAlaGlnGlyGlyTLeargeI 188  
326 AGCATGATTGGAAGCTTATCATGATATATACCAACAAGGACCGAGGG 575

seq\_name: /SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:Z98032

seq\_documentation\_block:

ID 298032 standard; cDNA; 1674 BP.

AC

DT

DE

KM

KW

KM

KW

KW

XX

XX

XX

XX

XX

PR

XX

XX

PI

XX

[illegible]









OM of: US-09-397-342-1 to: Issued\_Patents\_NA:\* out\_format : pfs  
Date: Apr 28, 2001 7:35 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPRO.spool/US09397342/runat\_27042001\_170000\_22911/app\_query.fasta.1.384  
-DB=Issued\_Patents\_NA -OPT=fastest -SOFT=p2n.rn1  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000  
-LOOEXT=0.000 -OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000  
-XGAEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -XGAPOP=10.000  
-YGAEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200  
-THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=200000000  
-USER=US09397342.ecgn1\_1\_59 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLFFX  
WAIT -THREADS=1

## Search information block:

Query: US-09-397-342-1  
Query length: 323  
Database: Issued\_Patents\_NA:\*  
Database sequences: 302621  
Database length: 87301344  
Search time (sec): 65.240000

## score list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-142-565-1 +		475.50	960.08	5.3e-47	1192
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-518-878B-38 +		472.50	973.07	1.3e-46	1255
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-294-522B-38 +		472.50	973.07	1.3e-46	1255
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-470-868A-38 +		472.50	973.07	1.3e-46	1255
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-807-861A-38 +		472.50	970.17	1.9e-46	1596
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-210-681-38 +		472.50	970.17	1.9e-46	1596
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-946-719A-38 +		472.50	970.17	1.9e-46	1596
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-937-456-3 +		472.50	967.76	2.6e-46	1949
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-172-528-3 +		472.50	967.76	2.6e-46	1949
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-318-199-3 +		472.50	957.08	1.0e-45	2782
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-937-466-1 +		469.50	957.08	1.0e-45	2782
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-172-528-1 +		469.50	957.08	1.0e-45	2782
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-318-199-1 +		469.50	957.08	1.0e-45	2782
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-518-878B-36 +		396.50	811.74	1.3e-37	1205
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-294-522B-37 +		396.50	811.74	1.3e-37	1205
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-807-861A-36 +		396.50	811.74	1.3e-37	1205
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-470-868A-36 +		396.50	811.74	1.3e-37	1205
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-210-681-36 +		396.50	811.74	1.3e-37	1205
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-946-719A-36 +		396.50	811.74	1.3e-37	1205
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-937-466-5 +		376.50	764.46	5.4e-35	1777
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-172-528-5 +		376.50	764.46	5.4e-35	1777
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-318-199-5 +		376.50	764.46	5.4e-35	1777
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-318-199-5 +		376.50	764.46	5.4e-35	1777
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-961-871-9 +		231.50	460.69	4.5e-18	1177
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-961-871-11 +		230.50	457.75	6.3e-18	1299
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-937-750C-61 +		227.50	448.51	2.1e-17	1594
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-234-613-61 +		227.50	448.51	2.1e-17	1594
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-933-750C-68 +		220.50	433.24	1.5e-16	1643
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-234-613-68 +		220.50	433.24	1.5e-16	1643
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-188-930-262 +		213.50	417.13	1.2e-15	1816
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-142-565-5 +		173.00	340.09	2.3e-11	847
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-142-565-3 +		173.00	249.79	2.5e-06	222
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-878B-14 +		123.00	245.80	4.2e-06	309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-294-522B-14 +		123.00	245.80	4.2e-06	309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-807-861A-14 +		123.00	245.80	4.2e-06	309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-470-868A-14 +		123.00	245.80	4.2e-06	309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-210-681-14 +		123.00	245.80	4.2e-06	309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-946-719A-14 +		123.00	245.80	4.2e-06	309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-946-719A-15 +		122.50	244.73	4.8e-06	309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-294-522B-15 +		122.50	244.73	4.8e-06	309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-807-861A-15 +		122.50	244.73	4.8e-06	309

seq\_name: /cgn2\_6/ptodata/1/lna/5A.COMB.seq:US-09-142-565-1  
seq\_documentation\_block:  
Sequence 1, Application US/09142565A  
Patent No. 6187560  
GENERAL INFORMATION:  
APPLICANT: Lee James Bealey  
APPLICANT: Kelly Paine  
APPLICANT: Robert James  
TITLE OF INVENTION: NOVEL COMPOUNDS  
FILE REFERENCE: GH-30002  
CURRENT APPLICATION NUMBER: US/09/142,565A  
CURRENT FILING DATE: 1999-06-30  
EARLIER APPLICATION NUMBER: 9704551.2  
EARLIER FILING DATE: 1997-03-05  
EARLIER APPLICATION NUMBER: 9705614.7  
EARLIER FILING DATE: 1997-03-18  
EARLIER APPLICATION NUMBER: 97305305.1  
EARLIER FILING DATE: 1997-07-16  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 1192  
TYPE: DNA  
ORGANISM: HOMO SAPIEN  
US-09-142-565-1

alignment\_scores:  
Quality: 475.50 Length: 316  
Ratio: 2.331 Gaps: 7  
Percent Similarity: 64.557 Percent Identity: 35.759  
alignment\_block:  
US-09-397-342-1 x US-09-142-565-1 ..  
Align seg 1/1 to: US-09-142-565-1 from: 1 to: 1192

20	AlaSerLysPheLeuLeuSerGlyCysAlaAlaThrValAlaGluLeuAl	36
238	GCTGTGAGTCTGCTGGGGGCGACGACGCTGTTTGTGCTGACCTGT	287
36	AThPheProLeuAspLeuThrLysThrArgLeuGluMetGlnGlyGlu	52
288	TACCTTTCACCTGACGACGACGACGACGCTGCTGACGACGACGACG	337
53	.....AlaAlaLeuAlaArgLeuGlyAspGlyAlaArgGluSerAla	66
338	ACGAGGGGCTGACGACGACGACGCTGCTGCTGCTGCTGCTGCTGCTG	366
67	ProTyrArgGlyMetValArgThrAlaLeuGlyIleIleGluGluGlu	83
367	CAGTACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	416
83	YPheLeuLysLeuThrPrgGlnGlyValThrProAlaIleTyrArgHisVal	100
417	TCCCTGACGACGCTGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG	466
100	AlTyrSerGlyArgMetValThrTyrGlnHisLeuArgGluValVal	116
467	GCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	516
117	PheGlyLysSerGlyAspGlnHisTyrProLeuThrLysSerValIleG1	133
517	ACCCCAAGGCGGGGAC...ACCTCCAGCTCCTACCTCCGAGATTGGC	563
133	YGLMetMetIaGlyValIleGlyGlnPheLeuAlaAsnProThrAspL	150

[illegible][illegible]

518 TTGCTCTGTAATGCCATGTGCACTGTGCTGAGCTGAGTGCACCTATGACCTC 567  
 214 ValLysHisTyrLeuValLeuAsnThrProLeuLuspaSniIleMeth 230  
 568 ATCAAGATGATCCCTCTGAAAGCCACCTCATGACAGATGACCTCCCTTG 617  
 230 rHisGlyLeuSerSerLeuCySergIleValAlaSerIleLeuGlyT 247  
 618 CCACTTCACTTCTGCTTGGGGGAGGCTTCTGACACCATGCTATCCGCT 667  
 247 hrProAlaSpValIleLysSerArgIleMetAsnGlnProArgAspLys 263  
 668 CCCCTGAGAGCTGTGTCAGACGAGATACATCACTGCCCTGGGCGCAG 717  
 264 GlnGlyArgGlyLeuLeuTyrLysSerSerThrAspGlyLeuIleGlnAl 280  
 718 .....TACAGTAGCGCTGGCCACCTGCTGCCCTTACCAT 749  
 280 aValGlnGlyGlyLeuGlyPheMetSerLeuTyrLysGlyPheLeuProSerT 297  
 750 GCTCCAGAGAGAGAGGCGCCGAGCCTTCTACAAAGGTTACATGCCCTCCT 799  
 297 rPleuArgMetThrProTyrSerMetValPheTyrPleuThrTyrGluLys 313  
 800 TTCCTCCCTGGGTCTCTGGAACGTGTGATGTCTGCACCTATGAGCAG 849  
 314 IleArg 315  
 850 CTGAAA 855

seq\_name: /cgn2\_6/ptodata/1/lna/5A\_COMB.seq:us-08-294-522B-38

seq\_documentation\_block:  
 ; Sequence 38, Application US/08294522B  
 ; Patent No. 5741666  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tartaglia, Louis A.  
 ; TITLE OF INVENTION: Compositions and Methods for the  
 ; TITILE OF INVENTION: Treatment of Body Weight Disorders, Including Obesity  
 ; NUMBER OF SEQUENCES: 48  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; ZIP: 10036-2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/294,522B  
 ; FILING DATE: 23-AUG-1994  
 ; CLASSIFICATION: 514  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 7853-015  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 790-9090  
 ; TELEFAX: (212) 869-8864/9741  
 ; INFORMATION FOR SEQ ID NO: 38:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1255 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-294-522B-38

alignment\_scores: Quality: 472.50 Length: 302

Ratio: 2.316 Gaps: 6  
 Percent Similarity: 67.550 Percent Identity: 34.437  
 alignment\_block:  
 US-09-397-342-1 x US-08-294-522B-38 ..

Align seq 1/1 to: US-08-294-522B-38 from: 1 to: 1255

18 ProArgAlaSer...LysPheLeuLeuSergIlySAlaAlaThrValAl 33  
 1 CTTACTGCCACTGTGAAGTTCTTGGGGCTGGCACACCTGCTTCATCCG 50  
 33 aGluLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuIleMetG 50  
 51 AGATCTATCACTTCTTCTGCTGATGCTGTAAGTCCGGTTACAGATCC 100  
 50 lngIlyGlnAlaAlaLeuAlaArgLeuGlyAspIly.....AlaArg 63  
 101 AAGAGAAAGT.....CAGGGGCCAGTGGCGGCTACA 132  
 64 GluSerAlaProTyrArgGlyMetValArgThrAlaLeuGlyIleIleG 80  
 133 GTCAGCGCCCGACGTACCGGGTGTATGGGCACCATCTGACCATGTGCG 182  
 80 uGlnGlyGlyPheLeuLysLeuTyrGlnGlyValIThrProAlaIleTyr 97  
 183 TACTGAGGGCCCCGAAAGCCTCTTACAAATGGGCTGTGCGGCTGCAGC 232  
 97 rGHisValAlaTyrSergIlyArgMetValIThrTyrGlnHisLeuArg 113  
 233 GCCAAATGAGCTTTGGCTTCCGTCATCGCCGCTGTATGATTCGTCAAA 282  
 114 GluValValPheGlyLysSergLusAspLusHisTyrProLeuTyrPlySe 130  
 283 CAG.....TTCTACACCAAGGCTCTGAGCATCCAGCATTTGGAGACCG 326  
 130 rValIleGlyGlyMetAlaGlyValIleGlyGlnPheLeuAlaAsp 147  
 327 CTTCTTGCAGGACGACACACAGGCTGCTGGCTGTGCTGTGCTGCGCAGC 376  
 147 rThrAspLeuValLysValGlnMetGlnMetGlyLysArgLysLeu 163  
 377 CCAGGATGTGTGAAGTCCGATTCAGAGTCCAGGCCG..... 417  
 164 GluGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaLysI 180  
 418 GCTGAGAGTGTGCGAGATACCAAGCACCTGATGCTTCAAGACCAT 467  
 180 eLeuAlaGlnGlyLysIleArgGlyLeuTyrPheAlaGlyTyrValProAsn 197  
 468 TGCCCCGAGAGAGAGGTTCCGGGCTCTTGGAAAGGACCTCTCCCATG 517  
 197 leGlnArgAlaAlaLeuValAsnMetGlyAspLeuThrTyrAspThr 213  
 518 TTGCTCGTAATGCCATGTGTCAACTGTGCTGAGTGGAGCTATGACCTC 567  
 214 ValLysHisTyrLeuValLeuAsnThrProLeuLuspaSniIleMeth 230  
 568 ATCAAGATGCTCTCTGAAAGCCACCTCATGACAGTGCATCCCTTG 617  
 230 rHisGlyLeuSerSerLeuCySergIleValAlaSerIleLeuGlyT 247  
 618 CCACTTCACTTCTGCTTGGGGGAGGCTTCTGACACCATGCTATCCGCT 667  
 247 hrProAlaSpValIleLysSerArgIleMetAsnGlnProArgAspLys 263  
 668 CCCCTGAGAGCTGTGTCAGACGAGATACATCACTGCCCTGGGCGCAG 717  
 264 GlnGlyArgGlyLeuLeuTyrLysSerSerThrAspGlyLeuIleGlnAl 280  
 718 .....TACAGTAGCGCTGGCCACCTGCTGCCCTTACCAT 749  
 280 aValGlnGlyGlyLeuGlyPheMetSerLeuTyrLysGlyPheLeuProSerT 297

```

750 GCTCCAGAGAGAGGGCCCGAGCCCTTCACAAAGGTTTCAGCCCTCCT 799
297 rpleuAargmethrProtrpSerMetValPheTrpleuThryrGluys 313
800 TTCTCCGCGTGGGTTCCGTGAACGTGTGATGTTCTGACCATGAGCAG 849
314 lilearg 315
850 CTCGAA 855

seq_name: /cgn2_6/ptodata/1/lna/5B_COMB.seq:US-08-470-868A-38

seq documentation block:
Sequence 38 Application US/08470868A
Patent No. 5861485
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis C.
TITLE OF INVENTION: Compositions and Methods for the
NUMBER OF INVENTION: Treatment of Body Weight Disorders, Including Obesity
CORRESPONDENCE ADDRESS:
ADDRESS: Pennile and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,868A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-0031-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-470-868A-38

alignment_scores:
Quality: 472.50 Length: 302
Ratio: 2.316 Gaps: 6
Percent Similarity: 67.550 Percent Identity: 34.437

alignment block:
US-09-397-342-1 x US-08-470-868A-38 ..

Align seg 1/1 to: US-08-470-868A-38 from: 1 to: 1255

18 ProAagAlaSer..LysPheLeuLeuSerGlycAlaAlaThrValAl 33
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1 CCTACTGCCACTGTGAGATTCTTGGGGGTGGCACAGCTGCTGCATCGC 50
33 aCluuAlaThrPheProLeuAspLeuThryrStHrArgLeuGlnMetg 50
||||| ||||||| ||||||| ||||||| ||||||| |||||||
51 ACATCTCATCACCTTTCCCTCTGTGATACGTCTAAAGTCCGCTTACGATCC 100
50 lnglgluAlaAlaLeuAlaArgLeuGlyAspGly.....AlaArg 63

```

101 AAGGAGAACT.....CAGGGCCCACTGGCCGTACA 132  
64 GluSerAlaProTyrArgGlyMetValArgThrAlaLeuGlyIleIleG1 80  
133 GTCAGCGCCCCAGTACCGGGGTGTGATGGCACCATTTGACCATGATGTCG 182  
80 uGluGluGlyPheLeuLeuLeuTyrPngLysIleValThrProAlaIleTyr 97  
183 TACTGAGGGGCCCCGAAGCCCTTCACAAATGGGCTGTTGGCCGCTGCAC 232  
97 rGhISaIValIATrSerGlyGlyArgMetValThrTyrGlnHisLeuArg 113  
233 GCCAAATAGACCTTTGGCTCTCTCCGATCGCGCCTGATGATTTGTCAAA 282  
114 GluValValPheGlyLysSerGluAspGlnHisTyrProLeuTyrPlysSe 130  
283 CAG.....TTCTACACCAAGGCGCTGAGATGCCAGCATTTGGAGCGCG 326  
130 rValIleGlyGlyMetIleAlaGlyValIleGlyGlnPheLeuAlaAsp 147  
327 CCTCTGTAGCAGGCGCAACACAGGCGCTGCTGGCTGGCGTGGCCAC 376  
147 rGhISaIValValIysValGlnMetGlnMetGluGlyLysArgLysLeu 163  
377 CCACGGATGTGGTAAAGCTCCGATTCAGCTCAGGCCGG..... 417  
164 GluGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaLysIle 180  
418 GCGGAGAGTGGTGGGAGATACCAAGACGCCATGCTTACAAAGACAT 467  
180 GluAlaGluGlyGlyIleArgGlyLeuTyrAlaGlyTyrValProAsnI 197  
468 TGGCCGAGGAAAGGGTTCGCGGGCTCTGGAAAGCAGCTTCCCAATG 517  
197 LeGIATrGAlaIaLeuValAsnMetGlyAspLeuThrThrTyrAspThr 213  
518 TTGCTGTAAATGCCATTGTCAACTGTGCTGAGCTGTGACCTATGACCTC 567  
214 ValIysHisTyrLeuValIleuAsnThrProLeuGlnIAspAsnIleMet 230  
568 ATCAAGAGATGCCCTCCGTAAGACCACTCATGACAGATGACTCCCTTG 617  
230 rHisGlyLeuSerSerLeuGlySerGlyLeuValAlaSerIleLeuGlyT 247  
618 CCACCTTCACCTTTCGCTTGGGGAGGCGCTTCGACACACTGTCATCGCCT 667  
247 hrProAlaAspValIleLysSerArgIleMetAsnGlnProArgAspLys 263  
668 CCCCTGTAGACGTGGTCAAGACGAGATCATGAACTCGCCCTGGGCCAG 717  
264 GluGlyArgGlyLeuLeuTyrLysSerSerThrAspGlyLeuIleGlnAl 280  
718 .....TACAGTAGCCCTGGCCACTGTGCCCTTACCT 749  
280 aValGlnGlyGluGlyPheMetSerLeuTyrLysGlyPheLeuProSerT 297  
750 GCTCCAGAAAGAGGGGCCCGAGGCTTCTACAAAGGTTACATGCGCTCT 799  
297 rPLeuArgMetThrProTyrSerMetValPheTyrPLeuThrTyrGluLys 313  
800 TTCTCGCTTGGGTTCTCTGMACTGTGGATGCTTCGACCATATAGACAG 849  
314 IleArg 315  
850 CTGAAA 855  
seq\_name: /cgn2\_6/ptodata/1/lna/5B\_COMB.seq:us-08-807-861A-38  
seq\_documentation\_block:  
; Sequence 38, Application us/08807861A  
; Patient No. 583975  
; GENERAL INFORMATION:

APPLICANT: Tartaglia, Louis A.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
REGULATION OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
NUMBER OF SEQUENCES: 64  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/807,861A  
FILING DATE: 26-FEB-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/518,878  
FILING DATE: 23-AUG-1995  
APPLICATION NUMBER: US 08/470,868  
FILING DATE: 06-JUN-1995  
APPLICATION NUMBER: US 08/294,522  
FILING DATE: 23-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-066  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1596 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-807-861A-38

alignment\_scores:  
Quality: 472.50 Length: 302  
Ratio: 2.316 Gaps: 6  
Percent Similarity: 67.550 Percent Identity: 34.437

alignment\_block:  
US-09-397-342-1 x US-08-807-861A-38 ..

Align seg 1/1 to: US-08-807-861A-38 from: 1 to: 1596

```
18 ProArgAlaSer...LysPheLeuLeuSerGlyCysAlaAlaThrValAl 33
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
343 CCTACTGACACTGTGAAGTTCTTGGGGGACACGCTGCGCTGACATCGC 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
33 agtLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetG 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
393 AGATCTCATACCTTCTCTGTGATACGTCAAGTCCGCTACAGATCC 442
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 lnglGluAlaAlaLeuAlaArgLeuGlyAspGly.....AlaArg 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443 AAGGCAAGAAAT.....CAGGGGCCAGTGGCGCGCTACA 474
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
64 GluSerAlaProTyrArgGlyMetValArgThrAlaLeuGlyIleIleG 80
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
475 GTCACGCGCCAGTACCGGGGTGTGATGGGACCATTCATGACCATGGTGC 524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
80 uGluGluGlyPheLeuLysLeuTrpGlnGlyValThrProAlaIleTyrA 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
525 TACTAGAGGCCCCGAGAGCCTCTACATAGGCTGTGGTGGCGGCTGACG 574
97 rghIsvaValTyrSerGlyArgMetValThrTyGlnHisLeuArg 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
575 GCCAATGAGCTTTCCTGCTGCCATCGGCGCTGATGATCTCTCAAA 624
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
114 GluValValPheGlyLysSerGlnAspGlnHisTyrProLeuTrpLys 130
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
625 CAG.....TTCTACACCAAGGCTGTGAGCATGCGACCATGGAGCGG 668
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
130 rValIleGlyGlyMetValAlaGlyValIleGlyGlnPheLeuAlaAsp 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
669 CCTCTAGAGGACACACACAGAGTCCCTGCTGTGGCTGTGGCGGCG 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147 rGthAspLeuValLysValGlnMetGlnMetGlnGlyLysArgLysLeu 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
719 CCACGATGTGTGAAGTCCATTCACAGGCTGAGCGCGG..... 759
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
164 GluGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaLys 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
760 GCTGAGGTGTGCGAGATCCAAAGCACCTCAATGCTTACAGACCAT 809
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
180 eLeuAlaGluGlyGlyLeuArgGlyLeuTrpAlaGlyTrpValProAsn 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
810 TCCCGGAGGAGGAGGTTCCGGGCTGTGAAAGGACCTTCCCAATG 859
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
197 leGlnArgAlaAlaLeuValAsnMetGlyAspLeuThrTyTrpAspThr 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
860 TTGCTGTGATGCTCATGTCACCTGTGAGCTGTGATGCTGATGACCTC 909
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
214 ValLysHisTyrLeuValLeuAsnThrProLeuGlnAspAsnIleMet 230
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
910 ATCAAGATGCTCTCTGAAACCACTCATGACATGACATGCTCTCTTG 959
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
230 rHisGlyLeuSerSerLeuGlySerGlyLeuValAlaSerIleLeuGly 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
960 CCACCTCACTTGTGCTTGGGAGGCTGTGACACCATGTGCATGCGCT 1009
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 hProAlaAspValIleLysSerArgIleMetAsnGlnProArgAspLys 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1010 CCCGTGACGCTGTGCAAGACGATGATGATGATGCTGCTGCGGCGCG 1059
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
264 GlnGlyArgGlyLeuLeuTyrLysSerSerThrAspCysLeuIleGln 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1060 .....TACAGTACGCTGCGCCACTGTGCTTACCAT 1091
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
280 aValGlnGlyGlyPheMetSerLeuTyrLysGlyPheLeuProSerT 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1092 GCTCCAGAGGAGGCGCCGAGCCTTCTACAAAGGTTTACATGCCCTCT 1141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
297 rPLeuArgMetThrProTrpSerMetValPheTrpLeuThrTyGlnLys 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1142 TTCCTCGGCTTGGGTCTGTGAGAGCGTGTGATGCTGCTACCTATGAG 1191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
314 IleArg 315
1192 CTGAAA 1197
```

seq\_name: /cgn2\_6/plodata/1/ina/6A\_COMB.seq:us-09-210-681-38

seq\_documentation\_block:  
Sequence 38 Application US/09210681  
Patent No. 6057109  
GENERAL INFORMATION:  
APPLICANT: Tartaglia, Louis A.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
REGULATION OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
NUMBER OF SEQUENCES: 64  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York

```

? COUNTRY: U.S.A.
? ZIP: 10036-2711
? COMPUTER READABLE FORM:
? MEDICAL TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/210,681
? FILING DATE:
? CLASSIFICATION:
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: 08/807,861
? FILING DATE: 26-FEB-1997
? APPLICATION NUMBER: US 08/518,878
? FILING DATE: 23-AUG-1995
? APPLICATION NUMBER: US 08/470,868
? FILING DATE: 06-JUN-1995
? APPLICATION NUMBER: US 08/294,522
? FILING DATE: 23-AUG-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Coruzzi, Laura A.
? REGISTRATION NUMBER: 30,742
? REFERENCE/DOCKET NUMBER: 7853-066
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 790-9090
? TELEFAX: (212) 869-9741/8864
? TELEX: 66141 PENNIE
? INFORMATION FOR SEQ ID NO: 38:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1596 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA
? US-09-210-681-38

alignment_scores:
? Quality: 472.50 Length: 302
? Ratio: 2.316 Gaps: 6
? Percent Similarity: 67.550 Percent Identity: 34.437

alignment_block:
US-09-397-342-1 x US-09-210-681-38 ..
Align seg 1/1 to: US-09-210-681-38 from: 1 to: 1596

18 ProArgLaser..LysPheLeuLeuSerGlyCysAlaAlaThrValAl 33
   ||| |||::: ||| ||| ||| ||| ||| ||| ||| ||| |||
343 CCTACTGCCACTGTGAAGTTCTTGGGGCTGGCACACCTGCTGCATCGC 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
33 agLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetg 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
393 AGATCTCATACCTTCTCTGCTGCTACTGCTAAAGTCCGGTTCAGATCC 442
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 InGlyAlaAlaLeuAlaArgLeuGlyAspGly.....AlaArg 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443 AAGGAGAAAGT.....CAGGGGCAAGTGGCGGCTACA 474
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
64 GluSerAlaProTyrArgGlyMetValArgThrAlaLeuGlyLeuIleG1 80
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
475 GTGAGGCGCCAGTACCGGGGTGATGGGACCATTCGTGACCATGGTGGC 524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
80 uGluGlyPheLeuLysLeuTyrGlnGlyValThrProAlaIleTyrA 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
525 TACTGAGGCGCCCGAAGCTCTACATATGGCTGGTGGCGGCTGACG 574
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
97 rGHisValValTyrSerGlyArgMetValArgThrTyrGlnHisLeuArg 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
575 GCCAAATGAGACTTGGCTGTCCGATCGGCTGATGATGATGTCGACAA 624
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
114 GluValValPheGlyLysSerGluAspGlnHisTyrProLeuTyrPlyse 130

```

```

625 CAG.....TTCTACACCAAGGCTGACCATGCCACATTCGGACCG 668
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
130 rValIleGlyGlyMetMetAlaGlyValIleGlyGlnPheLeuAlaAsp 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
669 CTCCTGACGAGGAGGACACACAGGTGCTGCTGCTGCTGCTGCTGCTG 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147 rGHisValValTyrSerGlyArgMetValArgThrTyrGlnHisLeu 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
719 CACGAGTGTGTAAAGTCCGATTCACAGCTCAGCGCCG..... 759
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
164 GluGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaVal 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
760 GCTGAGAGTGTGCGAGATACCAAGCACCGTCAATGCTTACAGACCAT 809
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
180 eLeuAlaGluGlyValLeuArgGlyLeuTyrPalaGlyTyrPalaPro 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
810 TGCCCGAGAGGAGGAGGTTCCGGGCTCTGGAAGGACCTTCCCATG 859
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
197 LeuArgAlaAlaLeuValAspMetGlyAspLeuThrThrTyrAspThr 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
860 TTGCTGCTAATGCCATGTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 909
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
214 ValHisHisTyrLeuValLeuAsnThrProLeuGluAspAsnIleMet 230
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
910 ATCAAGATGCTCCCTCTGGAAGCCACCATGATGACATGCTCCCTG 959
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
230 rHisGlyLeuSerSerLeuGlyLeuValAlaSerIleLeuGlyT 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
960 CCACCTTACTTCTGCTTGGGAGGCTTCTGACCATGCTGATGCTGCT 1009
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 hrProAlaAspValIleLysSerArgIleMetAsnGlnProArgAsp 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1010 CCCCTGACGAGGCTGACACGAGATACATGAACTGCTGCTGCTGCTG 1059
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
264 GluGlyArgGlyLeuLeuTyrLysSerThrAspCysLeuIleGlnAl 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1060 .....TACAGTAGGCTGGCCACTGCTGCTTACCAT 1091
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
280 aValGlnGlyGlyPheMetSerLeuTyrLysGlyPheLeuProSer 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1092 GCTCCAGAGAGGAGGCGCCGACCTTCTACAAAGGTTATGCTCCTCT 1141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
297 rPleuArgMetThrProTyrSerMetValPheThrPheThrTyrGly 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1142 TTCTCCGCTTGGCTCTGAGACGTGATGTTCTGACCTATGAGCAG 1191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
314 IleArg 315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1192 CTGAAA 1197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

seq_name: /cgn2_6/ptodata/1/lna/6B_COMB.seq:US-08-946-719A-38

seq_documentation_block:
? Sequence 38, Application US/08946719A
? Patent No. 6121017
? GENERAL INFORMATION:
? APPLICANT: Tartaglia, Louis A.
? TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
? TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
? NUMBER OF SEQUENCES: 64
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Pennie & Edmonds LLP
? STREET: 1155 Avenue of the Americas
? CITY: New York
? STATE: New York
? COUNTRY: U.S.A.
? ZIP: 10036-2711
? COMPUTER READABLE FORM:
? MEDICAL TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946,719A
FILING DATE: 8-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/807,861
FILING DATE: 26-FEB-1997
APPLICATION NUMBER: US 08/518,878
FILING DATE: 23-AUG-1995
APPLICATION NUMBER: US 08/470,868
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: US 08/294,522
FILING DATE: 23-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-107
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1596 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-946-719A-38

```

```

alignment_scores:
Quality: 472.50      Length: 302
Ratio: 2.316         Gaps: 6
Percent Similarity: 67.550      Percent Identity: 34.437

```

```

alignment_block:
US-09-397-342-1 x US-08-946-719A-38 ..

```

```

Align seg 1/1 to: US-08-946-719A-38 from: 1 to: 1596

```

```

18 ProatrgAlaser...LysPheLeuLeuSerGlyCysAlaAlaThVal1 33
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
343 CCTACTGCCACTGTGAAGTTCTTGGGCTGGCACAGCTGCTGCATGCGC 392
33 aglUleuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetG 50
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
393 AGATCTCATCACTTCTCTGATGATGCTAAAGTCCGGTTACAGATCC 442
50 lnglyglUAlaAlaLeuAlaArgLeuGlyAspGly.....AlaArg 63
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
443 AAGGGAAGT.....CAGGGCCAGTGGCGCCTACA 474
64 GluSerAlaProTyrArgGlyMetValArgThrAlaLeuGlyIleIleG1 80
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
475 GTCAGCGCCCACTACCGCGGTGTGATGGCACCATTTGACCATGTGCG 524
80 uGlnGluPheLeuLysLeuTrpGlnGlyValThrProAlaIleTyrA 97
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
525 TACTGAGGCGCCCGAGAGCTCTACATGAGGTGGTGGCGCCGCGAGC 574
97 rghsValValTyrSerGlyGlyArgMetValThrTyrGlnHisLeuArg 113
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
575 GCCAAATAGCTTGTGCTCTGTCCCATCGCGCTGTATGATTCGTCAAA 624
114 GluValAlaPheGlyLysSerGlnAspGlnHisTyrProLeuTrpLys 130
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
625 CAG.....TTCTACACCAAGGGCTGTAGCGTCCAGCATTTGGAGCGG 668
130 rValIleGlyGlyMetMetAlaGlyValIleGlyGlnPheLeuAlaAsp 147
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
669 CCTCTACAGGACACACAGGTGCGCTGTGCTGTGGTGTGGCCAGC 718

```

```

147 rothrasPleuValIysValGlnMetGlnMetGlnGlyLysArgLysLeu 163
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
719 CCAGGATGTGTGAAGGTCCGATTCAGAGCTCAGGCCG..... 759
164 GluGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaLysI1 180
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
760 GGTGAGGTGTGCGGATATACCAAGACCGCTCATGCTTCAAGACCAT 809
180 eleuAlaGlnGlyLysIleArgGlyLeuTrpAlaGlyTyrPalaProAsn1 197
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
810 TGCCGAGAGGAGAGGTTCGCGGCTCTGGAAAGGAGCTCTCCCAATG 859
197 leglnArgAlaAlaLeuValAsnMetGlyAspLeuThrThrTyrAspPhr 213
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
860 TTGCTCGTAATGCCATTTGTCAACTGTGCTGAGCTGTGATGACCTATG 909
214 ValIysHisTyrLeuValLeuAsnThrProLeuGlnAspAsnIleMetTh 230
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
910 ATCAAGATGGCCTCTCAAGCCAAAGCCACTCATGACATGACCTCCCTG 959
230 rHisGlyLeuSerSerLeuGlySerGlyLeuValAlaSerIleLeuGlyT 247
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
960 CCACCTCACTTCTGCTTGGGCGAGGCTCTGACCACTGTCAATCGCCT 1009
247 hrProAlaAspValIleLysSerArgIleMetValAsnGlnProArgAspLys 263
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1010 CCCCTGTAGAGTGTGTCAAGAGAGATATCATGACTGCGCTGGGCCAG 1039
264 GlnGlyArgGlyLeuLeuTyrLysSerSerThrAspCysLeuIleGlnAl 280
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1060 .....TACAGTGGCCTGGCCACTGCGCTTACCAT 1091
280 aValGlnGlyGlnGlyPheMetSerLeuTyrLysGlyPheLeuProSerT 297
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1092 GCTCCAGAGAGAGGCGCCCGAGCCTCTTCAAAAGGTTTCAATGCTCCT 1141
297 rPLeuArgMetThrProTrpSerMetValPheTrpLeuThrTyrGlnLys 313
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1142 TTCCTGCTTGGGTTCTGTGAACGTGTGATGTGTGTCACCATATGAGAG 1191
314 IleArg 315
|||||
1192 CTGAAA 1197

```

```

seq_name: /cgn2.6/ptodata1/ina/5B_COMB.seq:US-08-937-466-3

```

```

seq_documentation_block:

```

```

; Sequence 3, Application US/08937466

```

```

; Patent No. 5846779

```

```

; GENERAL INFORMATION:

```

```

; APPLICANT: Zhang, Ning

```

```

; APPLICANT: Amaral, M. Catherine

```

```

; TITLE OF INVENTION: UCP3 Genes

```

```

; NUMBER OF SEQUENCES: 6

```

```

; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

```

```

; STREET: 75 DENISE DRIVE

```

```

; CITY: HILLSBOROUGH

```

```

; STATE: CALIFORNIA

```

```

; COUNTRY: USA

```

```

; ZIP: 94010

```

```

; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: Patent in Release #1.0, Version #1.30

```

```

; CURRENT APPLICATION DATA:

```

```

; FILING DATE:

```

```

; CLASSIFICATION: 435

```

```

; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: OSMAN, RICHARD A

```

```

:
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: T97-009
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 343-4341
: TELEFAX: (650) 343-4342
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1949 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-937-466-3

```

```

alignment_scores:
  Quality: 472.50      Length: 300
  Ratio: 2.386         Gaps: 7
  Percent Similarity: 66.000  Percent Identity: 35.667

```

Alignment\_block:

US-09-397-342-1 x US-08-937-466-3

Align seg 1/1 to: US-08-937-466-3 from: 1 to: 1949

```

22  LysPheLeuLeuSerGlyCysAlaAlaThrValAlaGluLeuAlaThrph 38
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
285  AAGTTCCTGGGGCCGCGCAGCTGCTTTTGGGAGCCTCCCTCATT 334
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
38  eProlLeuAspLeuThrThrArgLeuGlnGlnGlnGlnGlnGlnGlnGln 55
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
335  TCCCTGGACACCGCCAGGTCGCTGACAGATCCAGGGGAG..... 377
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
55  eAlaIaArgLeuGlyAspGlyAlaArgGluSerAlaProTyrArgGlyMet 71
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
378  .....AACCCAGGGGCT...CAGAGCGTGCAGTACCCGCGTGTG 413
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
72  ValArgThrAlaLeuGlyIleIleGluGluGluGluGluGluGluGlu 88
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
414  CTGGGTACCTCTGATGATGTCGCGACAGAGGTCGCCGAGCCCTTA 463
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
88  pGlnGlyValThrProAlaIleTyrArgHisValValTyrSerGlyGly 105
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
464  CAGCGAGCTGTGCTGCTGCTGCGACGCCAGATGATTGCTCCCTCCAT 513
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
105  rGmetValThrTyrGlnHisLeuArgGluValValPheGlyLysSerGly 121
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
514  GAATTGGCCCTTAAGACTGTGTACAGAGTTTACACCCCGGAGGAGCG 563
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
122  AspGluHisTyrProLeuTyrLysSerValIleGlyGlyMetMetAla 138
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
564  GAC...CACTCCAGCGTCGCCATCAGGATTCGTGCGAGGTCGACAGAG 610
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
138  yValIleGlyIlePheLeuAlaAsnProThrAspLeuValValGln 155
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
611  AGCCATGAGGAGTGTGCTGCGCCAGCCAGAGTGTGTGAAGGTCCGAT 660
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
155  eGln.....MetGluLysArgLysLeuGly 165
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
661  TTCAGGCCATGATACGCTGGAACTGGAGAGAGAGAA..... 701
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
166  LysProLeuArgPheArgGlyValHisHisAlaPheAlaLysIleLeu 182
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
702  .....TACAGAGGAGTATGATGATGCTACAGAACCATCGCCAG 739
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
182  aGluGlyIleArgGlyLeuTyrPheAlaGlyTyrValProAsnIleGln 199
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
740  GAGGAGAGGAGTCAAGGCTGTGTGAAGAGGACTTGGCCCAACATCAAA 789
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
199  rGAlaIaLeuValAsnMetGlyAspLeuThrThrTyrAspThrValLys 215
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
790  GAATGCGCATGTTCACACTGTGCTGAGATGTGTGACCTACGACATCATCA 839
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

216  HisTyrLeuValLeuAsnThrProLeuGluAspAsnIleMetThrHis 232
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
840  GAGAACTTCTGGAGTGTCTACCTGTATTACGACAACTTCCCTGTACT 889
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
232  yLeuSerSerLeuCysSerGlyLeuValAlaSerIleLeuGlyThrPro 249
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
890  TGCTCTGCTTGGAGCTGTCTGTGTCCACAGTGTGCTCCCTCCCGG 939
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
249  LaAspValIleLeuSerArgIleMetAsnGlnProArgAspLysGlnGly 265
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
940  TGGATGTGTAAACACCGCATACATGAACTCC.....CTAGGC 980
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
266  ArgGlyLeuLeuTyrLysSerSerThrAspCysLeuIleGlnAlaVal 282
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
981  AGG.....TACGCGACCCCTCTGCTACATGTATGCTGAAGATGTGGC 1021
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
282  nGlyGluGlyPheMetSerLeuTyrLysGlyPheLeuProSerThrLeu 299
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1022  TCAGGAGGAGCCACAGCGCTTGTACAAAGATTTGTCCCTCTCTGTC 1071
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

seq\_name: /cgn2\_6/plodata/1/lna/5B\_COMB.seq:US-09-172-528-3

seq\_documentation\_block:

Sequence 3, Application US/09172528

Patent No. 5952469

GENERAL INFORMATION:

APPLICANT: Zhang, Ning

APPLICANT: Amaral, M. Catherine

TITLE OF INVENTION: UCP3 Genes

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 75 DENISE DRIVE

CITY: HILLSBOROUGH

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94010

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/172,528

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/937,466

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: T97-009

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 343-4341

TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1949 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-172-528-3

alignment\_scores: Quality: 472.50 Length: 300



Ratio: 2.386 Gaps: 7  
Percent Similarity: 66.000 Percent Identity: 35.667

alignment\_block:

US-09-397-342-1 x US-09-172-528-3 ..

Align seg 1/1 to: US-09-172-528-3 from: 1 to: 1949

```

22 LysPheLeuSerGlyCysAlaIleThrValAlaGluLeuAlaThrPh 38
   |||||||.....| ||||| |||||.....| |||||
285 AAGTTCCTGGGGGCGGCGACCTGCTGTTTGGGACCTCTCCACTT 334
   |||||||.....| ||||| |||||.....| |||||
38 eProLeuAspLeuThrIleThrArgLeuGlnMetGlnGlyAlaIle 55
   |||||||.....| ||||| |||||.....| |||||
335 TCCCTGGACACCGCAAGTCCGCTGCACATCCAAAGGGAG..... 377
   |||||||.....| ||||| |||||.....| |||||
55 euAlaIleArgLeuGlyAspGlyAlaIleArgGluSerAlaProTyrArgGlyMet 71
   |||||||.....| ||||| |||||.....| |||||
378 .....AACCAGGGCT...CAGACGCTGCACGTACCGCGCTGTG 413
   |||||||.....| ||||| |||||.....| |||||
72 ValArgThrAlaLeuGlyIleIleGlnGluGlnGlyPheLeuLysLeuTr 88
   |||||||.....| ||||| |||||.....| |||||
414 CTGGGTACCACTGCTACTGTGTGGCGACAGAGGGTCCCGACGCCCTA 463
   |||||||.....| ||||| |||||.....| |||||
88 pGlnGlyValThrProAlaIleTyrArgHisValValTyrSerGlyGly 105
   |||||||.....| ||||| |||||.....| |||||
464 CAGGGAGACTGGCTCGCTGGCTGCACCGCCAGATGATTGCTTCATTC 513
   |||||||.....| ||||| |||||.....| |||||
105 rGmetValThrTyrGlnHisLeuArgGluValAlaPheGlyLysSerGlu 121
   |||||||.....| ||||| |||||.....| |||||
514 GAATTGGCTCTAGACCTCTGCAAGCAGTTCTACACCCCAAGGAGACG 563
   |||||||.....| ||||| |||||.....| |||||
122 AspGlnHisTyrProLeuThrIleGlyLysMetMetAlaGlu 138
   |||||||.....| ||||| |||||.....| |||||
564 GAC...CACTCCAGCGTCGCATCAGGATCTGCGACGCTCCACGACG 610
   |||||||.....| ||||| |||||.....| |||||
138 YValIleGlyIlePheLeuAlaAsnProThrAspLeuValLysValGln 155
   |||||||.....| ||||| |||||.....| |||||
611 AGCCATGGCAGTGCACCTGCGCCACGCCAGGATGTGTGAAGGTCCAT 660
   |||||||.....| ||||| |||||.....| |||||
155 etGln.....MetGlnGlyLysArgLysLeuGlnGly 165
   |||||||.....| ||||| |||||.....| |||||
661 TTCAAGCCATGATACGCTGGGAACCTGGAGAGAGAGAA..... 701
   |||||||.....| ||||| |||||.....| |||||
166 LysProLeuArgPheArgGlyValHisIleAlaPheAlaLysIleLeuAl 182
   |||||||.....| ||||| |||||.....| |||||
702 .....TACAGAGGAGCTATGATGCTTACAGAACCATCGCCAG 739
   |||||||.....| ||||| |||||.....| |||||
182 aGlnGlyIleLeuArgGlyLeuThrPalaGlyTyrPalaProAsnIleGln 199
   |||||||.....| ||||| |||||.....| |||||
740 GGAGGAAGAGAGTCAAGGCTGTGGAAGGAGCTGGCCCAACATCACAA 789
   |||||||.....| ||||| |||||.....| |||||
199 rGAlaAlaLeuValAsnMetGlyAspLeuThrTyrAspThrValLys 215
   |||||||.....| ||||| |||||.....| |||||
790 GAAATGCCATGTCAACTGTGCTGATGATGCTACGACATCATCATCA 839
   |||||||.....| ||||| |||||.....| |||||
216 HisTyrLeuValLeuAsnThrProLeuGlnAspAsnIleMetThrHisG 232
   |||||||.....| ||||| |||||.....| |||||
840 GAGAACTTGGCTGAGCTTCACCTGTTACTACACACTTCCCTGTCTC 889
   |||||||.....| ||||| |||||.....| |||||
232 yLeuSerLeuLeuSerGlyLeuValAlaSerIleLeuGlnTyrProA 249
   |||||||.....| ||||| |||||.....| |||||
890 TGTCTCTCTTGGAGCTGGCTTGTGTCACAGTGGTGGCTCCCGG 939
   |||||||.....| ||||| |||||.....| |||||
249 IaAspValIleLysSerArgIleMetAsnGlnProArgAspLysGlnGly 265
   |||||||.....| ||||| |||||.....| |||||
940 TGAATGTGTAAGACCGCATCATGAACGCTCC.....CTAGGC 980
   |||||||.....| ||||| |||||.....| |||||
266 ArgGlyLeuLeuTyrLysSerSerThrAspCysLeuIleGlnAlaValG 282
   |||||||.....| ||||| |||||.....| |||||
981 AGG.....TACCGGACCCCTCTGCACCTATGCTGAAGATGTTGGC 1021
   |||||||.....| ||||| |||||.....| |||||
282 nGlyGlnGlyPheMetSerLeuTyrLysGlyPheLeuProSerThrLeu 299

```

```

1022 TCAGAGAGGACCCAGGCTCTCTACAAAGATTTGTGCTCTTCTGCG 1071
   |||||||.....| ||||| |||||.....| |||||
299 rGmetThrProTyrSerMetValPheThrPleuThrTyrGlnLysIleArg 315
   |||||||.....| ||||| |||||.....| |||||
1072 GTCTGGAGCTTGGAAAGTATGATGTTTGTACATATAGCAACACTGAG 1121

```

seq\_name: /cgn2\_6/plodata/1/lna/6A\_COMB.seq:US-09-318-199-3

seq\_documentation\_block:

```

? Sequence 3, Application US/09318199
? Patent No. 6025469
? GENERAL INFORMATION:
? APPLICANT: Zhang, Ning
? APPLICANT: Amaral, M. Catherine
? APPLICANT: Chen, Jin-Long
? TITLE OF INVENTION: UCP3 Genes
? NUMBER OF SEQUENCES: 6
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
? STREET: 75 DENISE DRIVE
? CITY: HILLSBOROUGH
? STATE: CALIFORNIA
? COUNTRY: USA
? ZIP: 94010
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/318,199
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/937,466
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: OSMAN, RICHARD A.
? REGISTRATION NUMBER: 36,627
? REFERENCE/DOCKET NUMBER: T97-009
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650) 343-4341
? TELEFAX: (650) 343-4342
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1949 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? US-09-318-199-3

```

alignment\_scores:

Quality	Ratio	Length
472.50	2.386	300

Percent Similarity: 66.000 Percent Identity: 35.667

alignment\_block:

US-09-397-342-1 x US-09-318-199-3 ..

Align seg 1/1 to: US-09-318-199-3 from: 1 to: 1949

```

22 LysPheLeuSerGlyCysAlaIleThrValAlaGluLeuAlaThrPh 38
   |||||||.....| ||||| |||||.....| |||||
285 AAGTTCCTGGGGGCGGCGACCTGCTGTTTGGGACCTCTCCACTT 334
   |||||||.....| ||||| |||||.....| |||||
38 eProLeuAspLeuThrIleThrArgLeuGlnMetGlnGlyAlaIle 55
   |||||||.....| ||||| |||||.....| |||||
335 TCCCTGGACACCGCAAGTCCGCTGCACATCCAAAGGGAG..... 377
   |||||||.....| ||||| |||||.....| |||||
55 euAlaIleArgLeuGlyAspGlyAlaIleArgGluSerAlaProTyrArgGlyMet 71

```

```

378 .....AACCCAGGGGCT...CAGAGCGTGCAGTCCCGGGTGTG 413
72 ValArgThrAlaLeuGlyIleIleGlyGlyGlyPheLeuLysLeuTr 88
414 CTGGGTACCATCTCTGACTATGTCGCGACAGAGGTCCCGCGACCCCTTA 463
88 pGlnGlyValThrProAlaIleTyrArgHisValValTyrSerGlyGly 105
464 CACGGGACTGTGGCTGGCTGCACCGCCAGATGATTTGGCTCCATTC 513
105 rgMetValThrTyrGlnHisLeuArgGlnValValPheGlyLysSerGln 121
514 GAATTTGGCTCTACGACTGTGTCAAGAGTTTACACCCCAAGGGAGCG 563
122 AspGlnHisTyrProLeuTrPlySerValIleGlyGlyMetMetAlaG 138
564 GAC...CACTCCAGCGTCCGCATCAGATTTCTGGACGGCTGCACGACG 610
138 yValIleGlyGlnPheLeuAlaAsnProThrAspLeuValValGlnM 155
611 AGCCATGGCAGTGCCTGCGCCAGCCAGATGTGGTGAAGTCCGAT 660
155 etcIn .....MetGlnGlyLysArgGlyLysLeuGly 165
661 TTCAAGCCATGATACGCTGGGAGACTGGAGAGAGAGAA..... 701
166 LysProLeuArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAl 182
702 .....TACAGAGGAGACTATGATGCTCCTACAGAACCATCCGCGAC 739
182 agLysGlyGlyIleArgGlyLeuTrPalaGlyTyrPalaProAsnIleGln 199
740 GGAGGAGAGAGTGCAGGGGCTGTGAAAGGAGCTTGGCCCAACATCAAA 789
199 rGAlaIleLeuValAlaAsnMetGlyAspLeuThrThrTyrAspThrLys 215
790 GAATATGCATTTGCACTGTGCTGAGATGGTGCATTCAGCATCATCAAG 839
216 HisTyrLeuValLeuAsnThrProLeuGlnAspAsnIleMetThrHisG 232
840 GAGAGTGTGCTGAGTCTCACCCTGTTTACGTACACACTCCCGCTGACT 889
232 yLeuSerSerLeuCySerGlyLeuValAlaSerIleLeuGlyThrPro 249
890 TGTCTGTGCTTGGAGCTGCTTCTGTCCACAGAGTGGTCCCTCCCGG 939
249 IaAspValIleLysSerArgIleMetAsnGlnProArgAspLysGlnGly 265
940 TGGATGTGGTAAAGACCCGATACATGACGCTCC.....CTAGGC 980
266 ArgGlyLeuLeuTyrLysSerSerThrAspCyLeuIleGlnAlaValG 282
981 AGG.....TACCGACGCCCTCGCACTGTATGCTGAAGATGGTGGC 1021
282 nGlyGlyGlyPheMetSerLeuTyrLysGlyPheLeuProSerTrpLeu 299
1022 TCAGGAGGAGCCACCGCCTTTCACAAAGATTTGGCCCTCTTCTCGC 1071
299 rgMetThrProTrpSerMetValPheTrpLeuThrTyrGlnLysIleArg 315
1072 GTCTGGGAGCTTGGACGTGATGATGTTGTAAATATGACGAACTGAAG 1121
seq_name: /cgn2_6/ptodata/1/1na/5B_COMB.seq:US-08-937-466-1
seq_documentation_block:
; Sequence 1, Application US/08937466
; Patent No. 5848779
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes

```

```

? NUMBER OF SEQUENCES: 6
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
? STREET: 75 DENISE DRIVE
? CITY: HILLSBOURGH
? STATE: CALIFORNIA
? COUNTRY: USA
? ZIP: 94010
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/937,466
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: OSMAN, RICHARD A
? REGISTRATION NUMBER: 36,627
? REFERENCE/DOCKET NUMBER: T97-009
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650) 343-4341
? TELEFAX: (650) 343-4342
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2782 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? US-08-937-466-1

```

```

alignment_scores:
Quality: 469.50 Length: 314
Ratio: 2.313 Gaps: 8
Percent Similarity: 64.650 Percent Identity: 34.713

```

alignment\_block:

US-09-397-342-1 x US-08-937-466-1 ..

Align seg 1/1 to: US-08-937-466-1 from: 1 to: 2782

```

22 LysPheLeuLeuSerGlyCysAlaAlaThrValAlaGluLeuAlaThrP 38
||||| ..... ||||| |||||
285 AACTTCTGTGGGGCGCGCACTGGCGCTGTTTGGCGACTCTCTCACTT 334
38 eProLeuAspLeuThrLysThrArgLeuGlnMetGlnGlyGlnAlaAla 55
||||| ..... ||||| |||||
335 TCCCTGTGACACCGCCCAAGGTCCGCTGTGCAGATCCAAAGGGAG..... 377
55 euAlaArgLeuGlyAspGlyAlaArgGluSerIleAlaProTyrArgGlyMet 71
378 .....AACCCAGGGGCT...CAGAGCGTGCAGTCCCGGGTGTG 413
72 ValArgThrAlaLeuGlyIleIleGlyGlyGlyPheLeuLysLeuTr 88
414 CTGGGTACCATCTCTGACTATGTCGCGACAGAGGTCCCGCGACCCCTTA 463
88 pGlnGlyValThrProAlaIleTyrArgHisValValTyrSerGlyGly 105
464 CACGGGACTGTGGCTGGCTGCACCGCCAGATGATTTGGCTCCATTC 513
105 rgMetValThrTyrGlnHisLeuArgGlnValValPheGlyLysSerGln 121
514 GAATTTGGCTCTACGACTGTGTCAAGAGTTTACACCCCAAGGGAGCG 563
122 AspGlnHisTyrProLeuTrPlySerValIleGlyGlyMetMetAlaG 138
564 GAC...CACTCCAGCGTCCGCATCAGATTTCTGGACGGCTGCACGACG 610
138 yValIleGlyGlnPheLeuAlaAsnProThrAspLeuValValGlnM 155

```



```
740 GGAGGAGAGTCAGGGCCCTGTGAAAGGAGCTGGCCCAACATCAACA 789
199 rgAlaIaIeUaValaSnMeGlyAspLeuThrThyAspThrValys 215
790 GAATGCGCATGTGCAACTGTGTGAGATGTGACCTACGACATCATCAAG 839
216 HisTyrLeuValLeuAsnThrProLeuGluAspAsnIleMetThrHisG 232
840 GAGAACTGCTGAGTCTCACCTGTTACTGACAACTTCCCTGTGCACTT 889
232 yLeuSerSerLeuCySerGlyLeuValAlaSerIleLeuGlyThrPro 249
890 TGCTCTGCTTGGAGCTGGCTTCTGTGACACAGTGTGGCTCCCGG 939
249 IaSpValIleLeuSerArgIleMetAsnGlnProArgAspLysGlnGly 265
940 TGGATGTGTGAAGACCCGATACATGACGCTCC.....CAGAGC 980
266 ArgGlyLeuLeuTyrLysSerSerThrAspCysLeuIleGlnAlaValG 282
981 AGG.....TACCGAGCCCTGTGCACTGTATGCTGAAGATGGCGGC 1021
282 ngLyGluGlyPheMetSerLeuTyrLysGlyPheLeuProSerThrLeu 299
1022 TCAGGAGGAGCCACGCGCTTCTACAAAGGATTGTGCCCTCTTCTGTC 1071
299 rgMetThrProTyrSerMetValPheThrLeuThyTyrGluLysIleArg 315
1072 GTCTGGAGAGCTTGAACGTGATGTTGTGTACATATGACACACGAGAG 1121
316 .....GluMetSerGlyValSerProPhe 323
1122 AGGCGCTTAATGAAGTCCAGTACGCGGGAATCTCCGTTT 1163

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-318-199-1
seq_documentation_block:
; Sequence 1, Application US/09318199
; Patent No. 6025469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/318.199
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

```
;
; LENGTH: 2782 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-09-318-199-1

alignment_scores:
      Quality: 469.50      Length: 314
      Ratio: 2.313      Gaps: 8
Percent Similarity: 64.650      Percent Identity: 34.713

alignment_block:
US-09-397-342-1 x US-09-318-199-1

Align seg 1/1 to: US-09-318-199-1 from: 1 to: 2782

22 LysPheLeuLeuSerGlyCysAlaAlaThrValAlaGluLeuAlaThrP 38
|||||
285 AAGTCTCTGGGGCGCGGCACTGCGGCTGTTTGGGAGCTCTCACTT 334
38 eProLeuAspLeuThrLysThrArgLeuGlnMetGlnGlyAlaAla 55
|||||
335 TCCCTGGACACCGCCAGAGTCCGCTGCGAGATCCAGAGGAG..... 377
55 euAlaArgLeuGlyAspGlyAlaArgGluSerAlaProTyrArgGlyMet 71
|||
378 .....AACCGAGGGCT...CAGAGCTGCGAGTACCGCGGTGTC 413
72 ValArgThrAlaLeuGlyIleIleGlnGluGlyPheLeuLysLeuThr 88
|||||
414 CTGGGTACCATCTCTACTATGTCGCGCACAGAGGTCCCGGAGCCCTTA 463
88 pgGlnGlyValThrProAlaIleTyrArgHisValValTyrSerGly 105
|||||
464 CACGGAGCTGTGCTGCGCTGCACCGCCAGATGATTGCTTCATTC 513
105 rgMetValThrTyrGlnHisLeuArgGluValAlaPheGlyLysSerGlu 121
|||||
514 GAATTGGCTCTACGACTCTGTCAAGCATTTACACCCCAAGGAGG 563
122 AspGlnHisTyrProLeuTyrLysSerValIleGlyMetMetAlaG 138
|||||
564 GAC...CACTCCAGCGCGGCATCAGATCTGCGAGGTCGACGACAG 610
138 yValIleGlyGlnPheLeuAlaAsnProThrAspLeuValLysGlnM 155
|||||
611 AGCCATGGCAGTGCCTGCGCCAGGCCACGGATGTGGTGAAGTCCGAT 660
155 etGln.....MetGluGlyLysArgLysLeuGluGly 165
|||||
661 TTCAGCCATGATACCGCTGGAGCTGGAGAGAGAGAA..... 701
166 LysProLeuArgPheArgGlyValHisHisAlaPheAlaLysIleLeu 182
|||||
702 .....TACAGAGGACTATGATGCTACAGAACATCTCCAG 739
182 agGlyGlyTleArgGlyLeuThrPalagLyrPvalProAsnIleGln 199
|||||
740 GGAGGAGAGTCAGGGCCCTGTGAAAGGAGCTGGCCCAACATCAACA 789
199 rgAlaIaIeUaValaSnMeGlyAspLeuThrThyAspThrValys 215
790 GAATGCGCATGTGCAACTGTGTGAGATGTGACCTACGACATCATCAAG 839
216 HisTyrLeuValLeuAsnThrProLeuGluAspAsnIleMetThrHisG 232
|||||
840 GAAAGTGTGTGAAGTCTACCTGTTACTGACAACTTCCCTGTGCACTT 889
232 yLeuSerSerLeuCySerGlyLeuValAlaSerIleLeuGlyThrPro 249
|||||
890 TGCTCTGCTTGGAGCTGGCTTCTGTGACACAGTGTGGCTCCCGG 939
```

```

249 laasphallllysserArglllemaSglnPrroargspylsgnngly 265
   ::::::::::::::::::::||| ||||||| |||
940 tggatgttgtaaaGCCGATACATGAAGCCTCC.....CTAGCG 580
266 ATGGTgLyLeuTyLysSerSerThrasPcYsLeuIlglnAlaValSl 282
   ||| ||:::|| |:::||||| :::
981 AG.....TACCGAGCCCCCTGCACGTGTATGCTAAGAATGGCGGC 1021
282 nclgyluglyPhemeSerLeuTyLysGlYPheIeuProserTyrPlea 299
   ||||| ::::||||| ||||| |||||
1022 TCAGGAGGACCACGCGCTTCACAAAGATTGTGGCCCTCTTCTCG 1071
299 rGmEthrProTrpserMetValPheTrpleuThrTyrgluIleay 315
   ||:::| ||:::||||| ||||| |||||
1072 GtCTGGAGACTTGTAACGTGATGATGTTGTTGACATATAGCAACTGAAG 1121
316 .....glumetSerGlyalSerProPhe 323
   ::::| ||||| |||
1122 AGGCGCTTAATGAAGTCCAGGATCGCGGATCTCCGTTT 1163

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-518-878B-36

seq_documentation_block:
: Sequence 36, Application US/08518878B
: Patent No. 5702802
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TITLE OF INVENTION: TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/518,878B
FILING DATE: 23-AUG-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-036
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8664
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1205 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-518-878B-36

alignment_scores:
Quality: 396.50 Length: 316
Ratio: 1.925 Gaps: 12
Percent Similarity: 65.190 Percent Identity: 33.228

alignment_block:
US-09-397-342-1 x US-08-518-878B-36 ..
Allsn seg 1/1 to: US-08-518-878B-36 from: 1 to: 1205
```

18 ProArgLaSer... LysPheLeuSerSerLysuysAlaIleThrValAl 33  
118 ProArgLaSer... LysPheLeuSerSerLysuysAlaIleThrValAl 33  
31 CCAACAGCCACGTGAAGTCTCTGGGGCTGGACACCTGCTGCATTCG 80  
33 agtLeuAlaIleThrPheProLeuAspLeuThrLysThrArgLeuGlnMetG 50  
113 agtLeuAlaIleThrPheProLeuAspLeuThrLysThrArgLeuGlnMetG 50  
81 AGATCTCATACACTTTCCTCCCTGGATACCGCAAGGTGCGGTGCAGATCC 130  
50 lngtLysLysLys... AlaLeuAlaArgLeuGlnLysAspGlyAlaArgLysSer 65  
113 lngtLysLysLys... AlaLeuAlaArgLeuGlnLysAspGlyAlaArgLysSer 65  
131 AAGGGAGAGGTCAAGGGCTAGTGGCC... ACCGACGACGAC 168  
66 AlaProTyrArgGlyMetValArgThrAlaLeuGlyLysIleIleGlnLysGly 82  
116 AlaProTyrArgGlyMetValArgThrAlaLeuGlyLysIleIleGlnLysGly 82  
169 GCCCATACCGGTGGCTTGGGTACACTCTTAACATGATGCGGCACTGA 218  
82 uGlyPheLeuLysLeuTyrPheGlnGlyValThrProAlaIleTyrArgHisV 99  
116 uGlyPheLeuLysLeuTyrPheGlnGlyValThrProAlaIleTyrArgHisV 99  
219 GGGTCCACGACACCTCTCAATAGAGGCTGTGGCGCGGCTGACGACCGCA 268  
99 aValAlaLysSerGlnGlyArgMetValIleThrArgHisLysLeuArgGlyVal 115  
269 TGAAGCTTGCTCCGCTCCGATATGGCTCTCAAGACTGTGTAAACAG... 315  
116 ValPheGlyLysSerGlnLysPheGlnHisThrProLeuTyrLysSerVal11 132  
316 ...TTCTACACCAAGGGCTCAAGACATGAGGACATCGGAGACCGCTCTCT 362  
132 eGlyGlyMetMetAlaGlyValIleGlyGlnPheLeuAlaAsnProThra 149  
363 GCGAGATACACCAAGGAGCGCTGGCGGTGTGTACCCAGCTACAG 412  
149 sPheuValLysValGlnMetGlnMetGlnGlyLysArgGlyLeuGlnGly 165  
413 ATGTGGTAAAGTCCGCTTCCACAGCTCCAGGCCG... GCTGGT 453  
166 LysProLeuArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAl 182  
454 GGTGGTCCGAGATACAGACACACTG... TCGAGCTACAAAGACATCAC 497  
182 agtLysGlyLysIleArgGlyLeuThrAlaGlyTyrValProAsnIleGlnA 199  
498 GAGAAGGAGGATCCGGGGCTCTGGAAAGGA... CTGCCCCAATGTGCC 544  
199 rgaAlaAlaLeuAlaAsnMetGlyAspLeuThrThrTyrAspThrValLys 215  
545 GTATGTCGATTCGATCACTGTCTGAGCTGGAGACCTATGACCTCATCA 594  
216 HisTyrLeuValLeuAsnThrProLeuGlnLysAsnIleMetThrHisGly 232  
595 GATACACTCT... CTGAGCCACTCATGACAAATGACCTCCCTGCACTT 641  
232 LysLeuSerLeuLysSerGlyLeuValAlaSerIleLeuGlnLysThrPro 249  
642 CACTTCTGCTTCGGGGCGGCTTGTGCACCACTGATGCCCTCCCTG 691  
249 LysAspValIleLysSerArgIleMetAsnGlnProArgAspLysGlnGly 265  
692 TGCATGTGGTCAGAGAGATACATGACT... 720  
266 ArgGlyLeuLeu... TyrLysSerSerThrAspCysLeuIleGlnAl 280  
721 ...CTGCTGGGCCAGTACACAGCGGCTACTGTGCCCTTACATG 764  
280 aValGlnGlyGlnGlyPheMetSerLeuTyrLysGlyPheLeuProSerT 297  
765 CTCGGAG... GAGGACCGCGGCTCTTCAACACAGGGGTATGCTTCT 811  
297 rPheuArgMetThrProTyrPserMetValPheThrPheThrTyrGlnLys 313  
112 TTTCTCGGCTTGGAGTCTTGAACAGTGAATGATGTTTGTACTTACCTTACAG 861









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2001, 16:17:21 ; Search time 1369.25 Seconds  
(without alignments)  
11190.817 Million cell updates/sec

Title: US-09-397-342-2

Perfect score: 1039  
Sequence: 1 ccgagctcgagcccgctat.....cagatatccatcacactgac 1039

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1283235 seqs, 7373929652 residues

Optimal number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

GenEmbl:\*

1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rtd:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_other:\*  
33: em\_htg\_rtd:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_vi:\*  
53: gd\_sts1:\*  
54: gd\_sts2:\*  
55: gd\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v11:\*  
59: gb\_v12:\*  
60: gb\_htg1:\*  
61: gb\_htg2:\*  
62: gb\_htg3:\*  
63: gb\_htg4:\*  
64: gb\_htg5:\*  
65: gb\_htg6:\*  
66: gb\_htg7:\*  
67: gb\_htg8:\*  
68: gb\_htg9:\*  
69: gb\_htg10:\*  
70: gb\_htg11:\*  
71: gb\_htg12:\*  
72: gb\_htg13:\*  
73: gb\_htg14:\*  
74: gb\_htg15:\*  
75: gb\_htg16:\*  
76: gb\_htg17:\*  
77: gb\_htg18:\*  
78: gb\_htg19:\*  
79: gb\_htg20:\*  
80: gb\_htg21:\*  
81: gb\_htg22:\*  
82: gb\_htg23:\*  
83: gb\_htg24:\*  
84: gb\_htg25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_ro1:\*  
95: gb\_ro2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	972	93.6	972	88	AF110532
2	785	75.6	2678	95	RNO300162
3	716.8	69.0	1890	95	RNO300164
4	715.2	68.8	2620	95	RNO300163
5	676.8	65.1	1726	9	AX061217
6	194.6	18.7	55820	81	AL512483
7	194.6	18.7	185614	60	AC008104
8	134	12.9	185614	60	AC008104
9	124	11.9	112126	12	AC007576
10	117.4	11.3	1525	12	AB024733
11	116.2	11.2	1434	15	ST0CPMRNA
					Y11220 Solanum tub



QY	880	gtctcaaggtgaagaggttcataagtcataaagaagcttttaaccatctgcgtgaagt	939
Db	841	gtttcaaggatgaaggatcattgactgattatataaagctttttaccatttggctgagaatg	900
QY	940	acccttgatcaatggtcttcgcgtactactatgaaaaatcaagagatgagtgaagtc	999
Db	901	acccttggtgcatatgctgttgcgcttacttataatgaaaaatcagagatgagtgaagtc	960
QY	1000	agtcacatttaa 1011	
Db	961	agtcacattttaa 972	
RESULT 2			
LOCUS	RNO300162	2678 bp	RNA
DEFINITION	Rattus norvegicus mRNA for uncoupling protein UCP-4 (Ucp-4 gene), isoform a.		
ACCESSION	RNO300162		
VERSION	AJ300162.1	GI:12055539	
KEYWORDS	isoform a; Ucp-4 gene; uncoupling protein UCP-4.		
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 2678)		
TITLE	Alberati-Giani, D., Gatti, S., Rial, E., Danel, F., Hauser, N., Bubendorf, C., and Barfai, T.		
JOURNAL	Three different isoforms of UCP-4 are expressed in rat preoptic anterior hypothalamus		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 2678)		
TITLE	Alberati-Giani, D.		
JOURNAL	Direct Submission		
FEATURES	Submitted (02-NOV-2000) Alberati-Giani D., Pharma Division		
source	Preclinical Research, CNS, F. Hoffmann-La Roche Ltd., Grenzacherstrasse 124, CH 4070 Basel, SWITZERLAND		
location/Qualifiers			
1..2678			
/organism="Rattus norvegicus"			
/strain="Sprague Dawley"			
/db_xref="taxon:10116"			
/sex="male"			
/tissue_type="brain, preoptic anterior hypothalamus"			
/dev_stage="6 weeks"			
/tissue_lib="UNI-ZAP custom cDNA library"			
/country="Switzerland"			
52..1020			
/gene="Ucp-4"			
52..1020			
/gene="Ucp-4"			
/function="uncouples respiration in eukaryotic cells"			
/codon_start=1			
/product="uncoupling protein UCP-4, isoform a"			
/protein_id="CAC20898.1"			
/db_xref="GI:12055540"			
/translation="MPPEESLOPFLNORMPRTSKFLISGCATVAELAPPPLDITKTRK			
RLNKGEGALAKIGDGMESAIVYSGMPTALGIVOEBSFLKNGVPAIVRHVYVSG			
GRMTVEHLREVRVYSGKSEDEHYPLMKSVTSGMAGVIGOTLANPTDLKVVOMQEKRR			
RLSEKPLRFGRVNHAFATLAEGGIRGTMAWMINIGDALVNNGDITTDVTRHHYLA			
LNATLEEDINATHGLSSLCGIVASTIGTPADIVTSRIINNPDRDQGRGLTKSSSTDCC			
IOAVEGEFLSLYKGLFPLSRMTPMVSEFWLITVEKIKQLSGVSPF"			
BASE COUNT	683 a	674 c	676 g 645 t
ORIGIN			
Query Match	75.6%;	Score 785;	DB 95; Length 2678;
Best Local Similarity	87.6%;	Pred. No. 3.3e-212;	
Matches 870; Conservative	0;	Mismatches 120; Indels 3; Gaps 1;	
20	tcgtctgcgtactacgtcgtatgctccgtcgcggagagagagagagagcttttgcgcgtcga	79	
1			

Db	32	TGCTCTGACCGTGGTGAATGTCCTTTCCC---TGAGGAGGAGATGCGGTACAGGCCATCTCA	88
QY	80	ccagagatgagcccccagcgagcaaatctcctaactgctccgctctgcgagctacacgttgccg	139
Db	89	CGCAGAGATGCGCCCGGACACCAAGTTCTCTACTGTCGGGCTGCCCGCACCGTGGCGCG	148
QY	140	agctggaaccttccccctgagatccacaaaactcgactcccaatgcaagaagaacgag	199
Db	149	ACGTAGCAACCTTTTCCCTCCGATCTCACAAAAACCCGACTCCAGATGCAAGGAGAACTCG	208
QY	200	ctctgtctcggtttggagacggtgcaagaagaaatcgtcccccatatagggaacggtgcga	259
Db	209	CCCTTGCTAAGTTGGGAGATGSGTGGCATGAGATCCGCCCTTACAGGGGCATGATGCGCA	268
QY	260	cagccctaggagatcatgtaaagagaaggcttctctaaagctttggcaagagtgagacccg	319
Db	269	CGGCGCTGGGAGTWTCCAGAGGAGAAAGGCTTCTTAAAGTTGTCGCAAGAGATGAGCGCCG	328
QY	320	ccatttacagaacgtagtgtaattctgtgaggttcgaatgtgcataatgaacatctccgag	379
Db	329	CCATTTCACGACACCTAGTACTCTGTGGAGGTCGGATGTCACCTACGAACTCCGCCGG	388
QY	380	aggtgtgtttggcaaaagtltgaagatltgaacatcatcccccttgggaataatgaafcatggag	439
Db	389	AAGTTGTGTTTGGCAAAAGTAGAATGACACATTATTCCTCTTGAGAGTCCGTCATTGGAG	448
QY	440	gagatgagctgtgtgtatltgtgcagatttllagccaatccaaactgaactagtgaaagtctc	499
Db	449	GGATATATGGCTTGGCTCATCGGACAGTTTTATGGCCAAACCCCATGACCTCGTGAAGGTCC	508
QY	500	agatgcaaatggaagaaaagaagaaactgtgaagagaacacatltcgattcgtgtgtac	559
Db	509	AGATCAAAATGGAAGGAAACCTTAACACTGGAAGGAAACCTTGAGATTCCTCGAGTAC	568
QY	560	atcatgcatcttgcacaaactctaagctgtgaaggagaaatgaagagcttggagagcttgg	619
Db	569	ATCATGCGGTTTCCAAAAATCTTAGCTGAAGAGGAATCCGTGGCTTGTGGCAGGCTGGA	628
QY	620	taccacaataaacaagaagcagcactggtgaaatatagtggagatltaaacactlataacag	679
Db	629	TACCCATATTTGAGAGAGCAGCAGCTTGTGAACATGGGAACTTTAACCATTTCAGACACAG	688
QY	680	tgaaacactacttgglatltgaatacaaccaacttgaaggacaatatcatgactcaacgtttat	739
Db	689	TGAAGCAGTACCTGGTATATGAACACAGCAGCTTGAAGACAAATATCGCACTCAGGCTTAT	748
QY	740	caagttatgtcttgagctggtgagcttctatctcttggaacaccacagccgaatgcatcaaaa	799
Db	749	CGAGTTGTGTTCTGGAGCTGTTGGCTTCTATTCTGGGAACACACAGCGGATGTCATCAAAA	808
QY	800	gcagaataatgatatcaaccacagagataaacaagaagaaggacttlytataaatcatcga	859
Db	809	GGCGAATATATGAACCAACCTCGAGACAAACAAGAAAGAGGGCTTTGTATTAAGTATCA	868
QY	860	ctgactgcttgatltcaagctgtltcaagtgtaagatltcaatgagctataaaagcgtttt	919
Db	869	CTGACTGCTGTATCCAGGCTGTTTCAAGGAGGAAGGCTTCTTGAGCCCTGTATTAAGGCTTTC	928
QY	920	taccatcttgagtgtaagagaccccttggtaatggtgttctgcttactacttgaaaaaa	979
Db	929	TACCTCTCCGTGTGGAATGACCCCTTGTCATATGTTGTGTGCTTACTTATGAAAAAA	988
QY	980	tcaagagatagtagtgagtcagtcacatttlaag 1012	
Db	989	TCAGACAGCTGAGTGGAGTCACTCCATTCTAAG 1021	

RESULT

3

RNO300164

1890 bp

RNA

LOCUS

Rattus norvegicus mRNA for uncoupling protein UCP-4 (Ucp-4 gene)

ISOFORM C.

DEFINITION

Accession

AJ300164

LOCUS

1890 bp

RNA

LOCUS

Rattus norvegicus mRNA for uncoupling protein UCP-4 (Ucp-4 gene)

ISOFORM C.

DEFINITION

Accession

AJ300164

LOCUS

1890 bp

RNA

LOCUS

Rattus norvegicus mRNA for uncoupling protein UCP-4 (Ucp-4 gene)

ISOFORM C.

DEFINITION

Accession

AJ300164

VERSION AJ300164.1 GI:12055543  
 KEYWORDS Isoform c; Ucp-4 gene; Uncoupling protein UCP-4.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 REFERENCE 1 (bases 1 to 1890)  
 AUTHORS Alberati-Giani D., Gatti S., Rial E., Danel F., Hauser N., Bubendorf C. and Barfal T.  
 TITLE Three different isoforms of UCP-4 are expressed in rat preoptic hypothalamus  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1890)  
 AUTHORS Alberati-Giani D.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-NOV-2000) Alberati-Giani D., Pharma Division Preclinical Research, CNS, F. Hoffmann-Lia Roche Ltd., Grenzachstrasse 124, CH 4070 Basel, SWITZERLAND  
 FEATURES  
 SOURCE Location/Qualifiers  
 1..1890  
 /organism="Rattus norvegicus"  
 /strain="Sprague Dawley"  
 /db\_xref="taxon:10116"  
 /sex="male"  
 /tissue\_type="brain, anterior preoptic hypothalamus"  
 /dev\_stage="6 weeks"  
 /tissue\_id="UNI-ZAP custom cDNA library"  
 /country="Switzerland"  
 160..1257  
 /gene="Ucp-4"  
 160..1257  
 /function="uncouples respiration in eukaryotic cells"  
 /codon\_start=1  
 /product="uncoupling protein UCP-4, isoform c"  
 /protein\_id="CAC20900.1"  
 /db\_xref="GI:12055544"  
 /translation="MPEPEESLOPLTQRMPTSKFLISGCAATVAELATPELDLTKT RLOMOGAALAKLDGAMESAPYRGMRTALGIQEGSLKMGVPAIYRHHVYSG GRMVTYHLREYVKGSEDEHYPLMKSVIGGMAGVIGIOPATLTKYOMOMEGKR RLEKPLRFRGVHNAFAKILAEGRIGRLMGWITNIOBALVYNGDLITDYVHYLY LNTALEDNIAHTHSSLSGSLVASILSTPADVILSRIMNDKDRQVGRGLYKSTDCV IQAVQEGFLSLYKGLPLSMRWKGTGRFLOFLYLSLCTLOVEFILLSDVMPHE RPIHLPEKNLYRCGFEPFOPFLRIMPELFD"  
 BASE COUNT 515 a 438 c 468 g 469 t  
 ORIGIN  
 Query Match 69.0%; Score 716.8; DB 95; Length 1890;  
 Best Local Similarity 87.0%; Pred. No. 8e-193;  
 Matches 800; Conservative 0; Mismatches 117; Indels 3; Gaps 1;

Oy 320 ccattacagacagtagtattcttggaggtcgaaatgagatgacacatccag 379  
 |||||  
 Db 437 CCATTACAGACAGCTAGTACTGTGAGAGTGGATGCTACCTACGACATCCGG 496  
 Oy 380 aggttggttggcaaaatgaaatgagatgacattatcccttggaaatcaatctgag 439  
 |||||  
 Db 497 AAGTGTGTTGGCAAAAGTGAAGTGAACATATATCCCTCTGGAAGTCCGTATTGGAG 556  
 Oy 440 ggaatgagctgtgtatcttggcagttcttagcccaatcccaatgagcttagtgaagctc 499  
 |||||  
 Db 557 GGATATGCTGGCTGCTATCGGACAGTATTAGCAACCCACCTACCTCGTGAAGTCC 616  
 Oy 500 agatgcaaatgaaagaaagaaacgtgaaagaaacatctgcatctgtgtgtgtac 559  
 |||||  
 Db 617 AGATCAATATGGAAGGAAGTGAAGTGAAGGAAGGAAGGAGTTCGTGAGTAC 676  
 Oy 560 atcaatgatttgcacaaatcttgaatgaaagaaatgagaggtcttggcagagctgtg 619  
 |||||  
 Db 677 ATCATGCTTTGCCAAAATCTTAGCTGAAGAGATCCGTGGCTTTGGCAGGCTTGA 736  
 Oy 620 taaccaatatacaagacagcaactggtgaaatgagatttaaccactatgatacag 679  
 |||||  
 Db 737 TACCAATATTCAGAGACAGCAGCTTGTGAACATGGAGATTTAACCATTACGACAG 796  
 Oy 680 tgaacactactgtgtatgtaatacaccaactgaaagaaatatacatgactcaggtttat 739  
 |||||  
 Db 797 TGAACCTACTACCTGTATTGAACACAGCAGTGAAGAACATATATGCCACTCAGGCTTAT 856  
 Oy 740 caagttatgttctgagactgtgacttctattcttggagaaacccagcagatcaatcaaa 799  
 |||||  
 Db 857 CCAGTTGTGTTCTGAGACTGGCTTCTTATCTTGGGAACACCCAGCATGTATCAACAAA 916  
 Oy 800 gaaagataatcaatcaacacagagataaagaaagaaagacttctgataatcatcga 859  
 |||||  
 Db 917 GCCGAATATGATCAACCACTCGAGCAACAAAGAAAGGCGCTTTGTATATAGTATCA 976  
 Oy 860 ctgactgttattcaaggctgttcaaggatgaagatcaatgactatataaagcttt 919  
 |||||  
 Db 977 CTGACTGTGATATCCAGCTGTTCAGAGAAAGGCTTCTGAGCTGTATTAAGCTTTC 1036  
 Oy 920 taccatctgctgaggaatg 939  
 |||||  
 Db 1037 TACCTCTGCTGCTGAGATG 1056  
 RESULT 4  
 RNO300163  
 LOCUS 2620 bp mRNA ROD 06-JAN-2001  
 DEFINITION Rattus norvegicus mRNA for uncoupling protein UCP-4 (Ucp-4 gene)  
 ACCESSION AJ300163  
 VERSION AJ300163.1 GI:12055541  
 KEYWORDS Isoform b; Ucp-4 gene; Uncoupling protein UCP-4.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 REFERENCE 1 (bases 1 to 2620)  
 AUTHORS Alberati-Giani D., Gatti S., Rial E., Danel F., Hauser N., Bubendorf C. and Barfal T.  
 TITLE Three different isoforms of UCP-4 are expressed in rat preoptic anterior hypothalamus  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 2620)  
 AUTHORS Alberati-Giani D.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-NOV-2000) Alberati-Giani D., Pharma Division Preclinical Research, CNS, F. Hoffmann-Lia Roche Ltd., Grenzachstrasse 124, CH 4070 Basel, SWITZERLAND  
 FEATURES  
 SOURCE Location/Qualifiers  
 1..2620  
 /organism="Rattus norvegicus"

BASE COUNT	657	a	661	c	679	g	623	t
ORIGIN	/strain="Sprague Dawley" /db_xref="taxon:10116" /sex="male" /tissue_type="brain, preoptic anterior hypothalamus" /dev_stage="6 weeks" /tissue_lib="UNI-ZAP custom cDNA library" /country="Switzerland" 139..1173 /gene="Ucp-4" 139..1173 /gene="Ucp-4" /function="uncouples respiration in eukaryotic cells" /codon_start=1 /product="uncoupling protein UCP-4, isoform b" /protein_id="CAC20899.1" /db_xref="GI:12055542" /translation="MPPPEESLOPIRNQWRPTSKFLISCAATVAATLAPPLDILKTR RLQNGERALAKLEDGAMESAPYRGMMRTALGIOECSFLKINGVTPALIRHYVYSGS GRWATYPLRFREVRHAKSEDEHYPLMKSVIGGMAGVIGQFLANPDTLVKQMOEGRR RLEKPPFRFGRVHAFKILAEGGIRGLMAGWIPRIQALVNNKGLDITDTVHYLV LNTALEDNIAHGLSCFLGSAVSIILCPAVISRLINQNRDQSGRLIYKSSDCCV IAAVQGGFLSLYKGLPSWLMRMDTLLHRCDLQGNATLVPSLETGLSDSSFLRLAI CLVOSHETNH"							

[illegible]

Db 716 TACCAGATATTCAGAGACAGACAGACTCTTGAGAACATGGGAGATTAAACCACTTACGACACAG 775

QY 680 tgaacactacttggatatttgaataacacacacttggagacaalatalcagactcogtllat 739

Db 776 TGAACCACTTACCTGGTATTTGAAACACAGGACCTTGGAAAGCAAAATATGCGCCACTCACCGCTTAT 835

QY 740 caagcttcatgtctcgcagcttgcagcttcctcatcttcggagacacacagccagatgtcataaa 799

Db 836 CCAGTTTGGTTCTGCGACTGGTGCGCTTATTCTTGGGAAACACAGCGCATGTCAATCAA 895

QY 800 gcaagaatatgaataacacacagagataaacaagaaggagacttltgtataatcatcoga 859

Db 896 GCCGATATATGACCAACCTCGAGCAAAACAAGGAAGGGCTTTTGATTAAGTCATCAAA 955

QY 860 ctgcactgtctgaatcagcgtgtccaagtggaagattcatatagctatataaagcttt 919

Db 956 CTGACTCGCGTATCCAGGCTGTTCAAAGGAAGAAGGCTTCCTGAGCGTGTATTAAGCTTTC 1015

QY 920 taccatcttgcgtggaatg 939

Db 1016 TACCTCTCTGGCTGAGAATG 1035

AX061217	AX061217	1726 bp	DNA	PAT	22-JAN-2001
LOCUS	Sequence	64 from Patent	W0078953.		
ACCESSION	AX061217				
VERSION	AX061217.1	GI:12406353			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
	1 (bases 1 to 1726)				
	Lai, P., Yang, J., Yue, H., Hillman, J. L., Tang, Y. T., Bandman, O.,				
	Burford, N., Baughn, M. R., Azimzai, Y., Lu, D. A., Au-Young, J. and				
	Patterson, C.				
TITLE	Human transport proteins				
JOURNAL	Patent: WO 0078953-A 64 28-DEC-2000;				
FEATURES	Incyte Genomics, Inc. (US)				
source	1. 1726				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	552 a 302 c 399 g 473 t				
ORIGIN					
Query Match	65.1%; Score 676.8; DB 9; Length 1726;				
Best Local Similarity	99.7%; Pred. No. 1.9e-181;				
Matches 678; Conservative	0; Mismatches 2; Indels 0; Gaps 0;				
332	acgtatgtatctctggaagtcgaatggtacacatacgaacatcccgagaagttgtttg 391				
452	acccagtatgtcttgagagtcgaatgctacatgtacaatcctccagagcttgcctttg 511				
392	gcaaaagtgaatgagcatatccctcttggaaatcaagtcatttgaggatgatgctg 451				
512	gcAAAAGTGAAGATGAGCATTTATCCCTTTGGAAATCACTGATTTGGAGGAGATGATGCTG 571				
452	ggttatctggtccagtttttaagccaatccaactgacactgaatgaagttcgaatgcaaatg 511				
572	gttttatgttggcagtttttagcccaatccaactgacactgaatgaagttcgaatgcaaatg 631				
512	aaggaaaaaagaactctgagaagaaaaccattgcgatttcgttgtacatcgtcatgtt 571				
632	aaggAAAAAAGAAACGAGAGGAAAAACCACTTGCATTTGCTGCTACATCATGCAATTTG 691				
572	caaaactcttaagctgaaggagaatacgaaggcttggcgaagctggtgatcccaatatac 631				
692	caAAATATCTTACTCTGAAGAGGAATAACGAGGGCTTTGGCCAGCTGGGTATCCCAATATAC 751				



http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR

```

/chr=18"

```

```

/misc_feature /map="18"
/misc_feature /clone="RP11-446F17"
/misc_feature 1. .1109
/misc_feature /note="assembly_fragment"
/misc_feature 1210. .3242
/misc_feature /note="assembly_fragment"
/misc_feature 3343. .4984
/misc_feature /note="assembly_fragment"
/misc_feature 5085. .8149
/misc_feature /note="assembly_fragment"
/misc_feature 8250. .10485
/misc_feature /note="assembly_fragment"
/misc_feature 10586. .12579
/misc_feature /note="assembly_fragment"
/misc_feature clone_end:SP6
/misc_feature vector_side:right"
/misc_feature 12680. .15104
/misc_feature /note="assembly_fragment"
/misc_feature 15205. .19613
/misc_feature /note="assembly_fragment"
/misc_feature 19714. .23297
/misc_feature /note="assembly_fragment"
/misc_feature 23398. .27445
/misc_feature /note="assembly_fragment"
/misc_feature 27346. .33569
/misc_feature /note="assembly_fragment"
/misc_feature 33670. .39886
/misc_feature /note="assembly_fragment"
/misc_feature 39987. .46130
/misc_feature /note="assembly_fragment"
/misc_feature clone_end:T7
/misc_feature vector_side:right"
/misc_feature 46231. .53403
/misc_feature /note="assembly_fragment"
/misc_feature 53504. .61709
/misc_feature /note="assembly_fragment"
/misc_feature 61810. .70390
/misc_feature /note="assembly_fragment"
/misc_feature 70491. .83671
/misc_feature /note="assembly_fragment"
/misc_feature 83772. .99035
/misc_feature /note="assembly_fragment"
/misc_feature 99136. .114877
/misc_feature /note="assembly_fragment"
/misc_feature 114978. .132866
/misc_feature /note="assembly_fragment"
/misc_feature 132967. .157662
/misc_feature /note="assembly_fragment"
/misc_feature 157763. .185614
/misc_feature /note="assembly_fragment"
BASE COUNT 56535 a 35534 c 35072 g 56364 t 2109 others
ORIGIN
Query Match 18.7%, Score 194.6, DB 60, Length 185614;
Best Local Similarity 93.5%; Pred. No. 4.1e-44;
Matches 203; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 121 tgcgagctacgctgctgagctagcaacattccctggagctacacaaactgacac 180
DB 82992 TTCCCTGATTTGTTGTTTGTATGCAACCTTTCCCTGATCTCACAACCTGACTC 82933
QY 181 caaatgcgaaggaagcagctctctcgttgagagcagtgcaagaagaatcgtccccc 240
DB 82932 CAAATGCAAGGAGAGACACTTTCTGCTGCGGTTCGAGAGAGGTGCAAGAACTGCCCC 82873
QY 241 tatagggaatgctgagcagccttaggagatcatggaaggaagagcttcttaagctt 300
DB 82872 TATAGGGAATGCTGCGACACTCTAGGAGATCATTTGAAGAGAGAGGCTTTAAAGCTT 82813
QY 301 tggcagaagtgagaccccgccattacagacagctag 337

```

```

DB 82812 TGGCAAGAGTGACACCCGCAATTTCAGACACGTAG 82776
RESULT 8
AC008104
LOCUS AC008104 185614 bp DNA HTG 26-MAY-2000
DEFINITION Homo sapiens chromosome 18 clone RP11-446F17 map 18, WORKING DRAFT
SEQUENCE, 22 unordered pieces.
ACCESSION AC008104
VERSION AC008104.3 GI:8072415
KEYWORDS HTG, HTGS, PHASEL, HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 185614)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 18, clone RP11-446F17
Unpublished
2 (bases 1 to 185614)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckery,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeCarliano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferrelita,P., Fitzhugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Lien,C., Locke,K., MacDonald,P.,
Markus,N., McEwan,P., McGurk,A., McKenna,K., McLaughlin,J.,
Meldrum,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (22-JUL-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7321473.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIDR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1022
Center clone name: 446_F_17
----- Summary Statistics
Sequencing vector: M13; M77815; 92% of reads
Sequencing vector: Plasmid; n/a; %-0.1% of reads
7.56620428751576Chemistry: dye-primer-merism; 92% of reads
Chemistry: dye-terminator Big Dye; 8% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 166796 bases at least Q40
Consensus quality: 174175 bases at least Q30
Consensus quality: 177996 bases at least Q20
Insert size: 19200; agarose-fp
Insert size: 18514; sum-of-ctrls
Quality covera.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1109: contig of 1109 bp in length
* 1110 1209: gap of 100 bp
* 1210 3242: contig of 2033 bp in length

```



*	3243	3342:	gap of	100	bp	in	length
*	3343	4984:	contig of 1642	bp	in	length	
*	4985	5084:	gap of	100	bp		
*	5085	8149:	contig of 3065	bp	in	length	
*	8150	8249:	gap of	100	bp		
*	8250	10485:	contig of 2236	bp	in	length	
*	10486	10585:	gap of	100	bp		
*	10586	12579:	contig of 1994	bp	in	length	
*	12580	12679:	gap of	100	bp		
*	12680	15104:	contig of 2425	bp	in	length	
*	15105	15204:	gap of	100	bp		
*	15205	19613:	contig of 4409	bp	in	length	
*	19614	19713:	gap of	100	bp		
*	19714	23297:	contig of 3584	bp	in	length	
*	23298	23397:	gap of	100	bp		
*	23398	27445:	contig of 4048	bp	in	length	
*	27446	27545:	gap of	100	bp		
*	27546	33569:	contig of 6024	bp	in	length	
*	33570	33669:	gap of	100	bp		
*	33670	39886:	contig of 6217	bp	in	length	
*	39887	39986:	gap of	100	bp		
*	39987	46130:	contig of 6144	bp	in	length	
*	46131	46230:	gap of	100	bp		
*	46231	53403:	contig of 7173	bp	in	length	
*	53404	53503:	gap of	100	bp		
*	53504	61709:	contig of 8206	bp	in	length	
*	61710	61809:	gap of	100	bp		
*	61810	70390:	contig of 8581	bp	in	length	
*	70391	70490:	gap of	100	bp		
*	70491	83671:	contig of 1181	bp	in	length	
*	83672	83771:	gap of	100	bp		
*	83772	99035:	contig of 15264	bp	in	length	
*	99036	99135:	gap of	100	bp		
*	99136	114677:	contig of 15742	bp	in	length	
*	114878	114977:	gap of	100	bp		
*	114978	132866:	contig of 17889	bp	in	length	
*	132867	132966:	gap of	100	bp		
*	132967	157662:	contig of 24696	bp	in	length	
*	157663	157762:	gap of	100	bp		
*	157763	185614:	contig of 27852	bp	in	length	

```

ores
source
Location/Qualifiers
1..185614
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-446F17"
/clone_1fb="RPCI-11 Human Male BAC
1..1109
/note="assembly_fragment"
1210..3242
/note="assembly_fragment"
3343..4984
/note="assembly_fragment"
5085..8149
/note="assembly_fragment"
8250..10485
/note="assembly_fragment"
10586..112579
/note="assembly_fragment
clone_end:sp6
vector_side:right"
12680..15104
/note="assembly_fragment"
15205..19613
/note="assembly_fragment"
19714..23297
/note="assembly_fragment"
23398..27445
/note="assembly_fragment"
27546..33569
/note="assembly_fragment"
33670..39886
```

```
misc_feature      /note="assembly_fragment"
33987. .46130
/note="assembly_fragment
clone_end:r7
vector_side:right"
46231. .53403
/note="assembly_fragment"
53504. .61709
/note="assembly_fragment"
61810. .70390
/note="assembly_fragment"
70491. .83671
/note="assembly_fragment"
83772. .99035
/note="assembly_fragment"
99136. .114877
/note="assembly_fragment"
114978. .132866
/note="assembly_fragment"
132967. .157662
/note="assembly_fragment"
157763. .185614
/note="assembly_fragment"
56535 a 35534 c 35072 g 56364 t 2109 others
ORIGIN
```

Query Match	12.9%	Score 134	DB 60	Length 185614
Best Local Similarity	96.5%	Pred. No. 7.1e+27		
Matches 137	Conservative	0	Mismatches 5	Indels 0
				Gaps 0

QY 9 ggaaccggttcctcgtctgcgctaactgtaatgtcccgtagagagaagagct .68  
||| ||||||||||||||||||||||||||||||||||||||||||||  
Db 50849 GGAGTGGCTATTCGTCCTTGCGCACTGCATGATCCGTCCCGAGAGGAGAGAGGCT .50908

Oy 69 ttgcgcgtgacccagagatggccccgaqcgacaattcccactgtccgcgtcgcgc 128  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 50909 TTGGCGGTGACCCAGAGATGCCCGGAGGAGCAATTCTACTGTCCGGCTGC GGSC 50968

QY 129 tacctgtgccgagctagcaacc 150  
|||||  
Db 50969 TACCGTGGCCGAGCTAGGTACC 50990

RESULT	9
AC007576/c	
LOCUS	AC007576 112126 bp DNA
DEFINITION	Arabidopsis thaliana chromosome I BAC F7A19 genomic sequence, complete sequence.

VERSION	AC007576.3	GI:5019264
KEYWORDS	HTG.	
SOURCE	thale cress.	

ORGANISM  
*Arabidopsis thaliana*  
 Eukaryota: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 112126)  
 REFERENCE  
 Federspiel, N. A., Palm, C. J., Conway, A. B., Conn, L., Hansen, N. F.,  
 AUTHORS

JOURNAL  
REFERENCE

Altafi, H., Araujo, R., Hultzer, L., Rowley, D., Buehler, E., Dunn, P., Gonzalez, A., Kremetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S., Luros, S., Schwartz, J., Shinn, P., Tortumli, M., Vysotskaia, V. S., Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R. W.  
Unpublished  
2 (bases 1 to 112126)

**AUTHORS**  
Fedeerspigel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, E.N., F.  
Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,  
Gonzalez, A., Kremenetskaia, I., Kim, C., Ienz, C., Li, J., Liu, S.,  
Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Yokotsuka, V.,  
Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.

**TITLE**  
Direct Submission

**JOURNAL**  
Submitted (17-MAY-1999) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
USA  
3 (bases 1 to 112126)  
Federpiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Araujo,R., Huizar,I., Rowley,D., Buehler,E., Dunn,P.,  
Gonzalez,A., Kremetskaja,I., Kim,C., Lenz,C., Li,J., Liu,S.,  
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaja,V.,  
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
Direct Submission  
Submitted (08-JUN-1999) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
4 (bases 1 to 112126)  
Federpiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Araujo,R., Huizar,I., Rowley,D., Buehler,E., Dunn,P.,  
Gonzalez,A., Kremetskaja,I., Kim,C., Lenz,C., Li,J., Liu,S.,  
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaja,V.,  
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
Direct Submission  
Submitted (16-JUN-1999) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
On Jun 8, 1999 this sequence version replaced gi:4887737.  
e-mail for correspondence: arabes@sequence.stanford.edu  
Genes with similarity to proteins in the databases are described as  
'putative', '-like' or 'similar to'. Genes that have EST  
similarity but no significant protein similarity are described as  
'unknown proteins'. Genes that are annotated based only on gene  
prediction software are described as 'hypothetical proteins'.  
The software programs used to predict genes include: Grail  
(Informatics Group, Oak Ridge National Laboratory,  
http://compbio.ornl.gov/section/index.html), GENSCAN (Chris Burge,  
http://genome.stanford.edu/~chris/GENSCAN.html), Fexa (V.Solovay  
& A.Salanov, Sanger Centre, http://genomic.sanger.ac.uk/),  
NetPlantene (S.M. Hebsgaard, et al., CBS, Technical University of  
Denmark, http://www.cbs.dtu.dk/NetPlantene.html) and  
EMOTif (Newill-Manning, C.G. Wu, T.D. & Brutlag, D.L.,  
http://motif.stanford.edu/projects.html).  
Location/Qualifiers  
1..112126  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/clone="F7A19"  
/chromosome="1"  
/map="92358"  
1240..2806  
/gene="F7A19.1"  
join(1240..1265,1351..1448,1540..1697,1788..1887,  
1976..2354,2426..2507,2612..2806)  
/gene="F7A19.1"  
/note="Hypothetical protein"  
/codon\_start=1  
/protein\_id="AAd39278.1"  
/db\_xref="GI:5080768"  
/translation="MDRLIKOTRRKHRAPOEKIREVSSSTREKKVARKSVSEKEDK  
KPSNMLOKFSKQMSQSTYPIGEMDAVAATATAIAIEETLWLNTHSGELGGS  
SSSKSRSEELPLEEPRLSRKSGSLIDSTKDHKPTLSPARKSSVAKITS  
MNLMDGHTKONODEEHEROKRPVSEPPRIOPPLRSEPPRAPPPPLSPSSPRL  
LPRETRROSESEHSRKDSDTADAMEKAEISKIAREYENLRKIDLEWAKREKARRK  
LDISESELEORRRKRGLOFRREDTEYIEQIAAGRAQAEDKROSKKEKVKACVITS  
TKLPGNACPF"  
4106..4769  
/gene="F7A19.2"  
4106..4769  
/gene="F7A19.2"  
4143..4610  
/gene="F7A19.2"  
/note="Unknown protein; location of ESTs 128A14T7,  
gb|U44625 and 128A14XP 3', gb|AI099841"  
/codon\_start=1  
/protein\_id="AAd39279.1"  
/db\_xref="GI:5080769"  
/translation="MNFISDQVKKLSSSTPEPDHKNKPEVETATAPATNAELMASA  
KVAEAQAARNESEKLDKGVAGASADILDAEAKYCKDEKSSGQYIDAKREKTLN  
DYESSHSTGAGGPPPTSOAEPASQPPAPKADDESGGLGGYAKMAQGLK"  
complement(5755..8975)  
/gene="F7A19.3"  
complement(join(5755..7820,7930..8641,8781..8975))  
/gene="F7A19.3"  
/note="Hypothetical protein"  
/codon\_start=1  
/protein\_id="AAd39280.1"  
/db\_xref="GI:5080770"  
/translation="MGTKVHCESLFGYHHSMGLANKESNNGCRMPLEFGYDNTSASN  
NDCCYNGPSTQTFGEFDKDYELIHHVYETQKDMDELKRPQNKEMQTEISLSO  
ATNDVYKMKIPSPPLANSYDRPSMSYVDNGHSPKSGNSGPPVSWGASKSVE  
VSEVPRKIRKMKIDLCIPADEYIDDEVEYELKDHVCSQILPBGVKTSSRIDG  
VRIGGSSRNGADLNEPDAOEANGAHSIDNGEFOGHRIDPBGVKSNGSVR  
EHIPVILPDGLDNKPKVMSHOPJRTYDHTGKASAPFOPAPKPPDLSLQPMVLN  
SSORVMSLPSNGSPSKAVVWRBERFIDLEADTONSHAEVITHONHDSLSLHOOP  
LYPIYRDSAVSNHLLSSNQNSFPGQVASAKRPVYLMSDTLGNKQKGYLGD  
RLQESNRFSDSGGNSRLNHNETWETSCKSKVAGTGYTPNGGRSDYSERKE  
VRDLNLTWLTNSVYEVKDEHLATLPLMLNPKVCSNSELADGKWSLSDAVPS  
SLKPLDIRDEAGDKVONIMWLERLKSQCSNNPVTEKIDANKETPGIAYDQSNVGA  
KVHYRMLDMNEPCDPLMDDQOETQETERYVASNCOIDLNLGSDDEGNCVPA  
SSRLNSKAPMIDLETVEPDESDDEEDGNISEKSEVKSIEKTEKPEPEKLAETI  
VAIYSAQIDREVEVYVASSVSETILLMWEAEYNTHEKNDLKDPTFSRQNSIEDI  
DYRESATLQLPDISDEYMRKPLPEMLKLEETGTALVTSQRONARKGQSRDF  
QRIIDLPGILSLSHVEYEDIQMDGFEMRAKRWMTPTGLTKRTGSGRRRAITTPA  
EPVYCPVAPAPAPPSPVQOYHVSQNSNTGEMCLEDRSFAGMGKMRPRRRCPPSAS  
TTTATTSNQSSSHAPPT"  
16683..17720  
/gene="F7A19.4"  
/note="Eif-5A"  
join(16683..16802,17074..17196,17290..17346,17420..17494,  
17619..17720)  
/gene="F7A19.4"  
/note="Eif-5A"  
/codon\_start=1  
/product="initiation factor 5A-4"  
/protein\_id="AAd39281.1"  
/db\_xref="GI:5080771"  
/translation="MSDEHHFESSDAGASKTYPOQAGTRKNGYIYIKRPPCKVVEY  
STSKTGKHAKCFVADIDFTKSKLDIYSSHNCVPIVNTDQIDLISDGYVS  
LITDNGSTKDKDILPNDDTLQDKSFDGKRLVYSVMSAMEEQINAKDIGNPK"  
19674..22095  
/gene="F7A19.5"  
join(19674..20007,20763..21468,21551..21715,21837..22095)  
/gene="F7A19.5"  
/note="Similar to DNA-binding proteins"  
/codon\_start=1  
/protein\_id="AAd39282.1"  
/db\_xref="GI:5080772"  
/translation="MSEKEEAPSTSKSTGASRPRLSLPRPSEMEFNGVGFSPRP  
MTLVNMFPPDSDEFERSFQLAGAMSPATAAATAASDQRLDEGTNSSGDVD  
PRFKONPRVAVLDICNIYOGSYGMFTLOAOVAYANMMOPOETPEPSSOYQ  
SPSSGOOIPITSAIPAPORETSQVITIEHSSOOLNADKADGCMYRKQYGVGS  
EPRRSYKCTNPGCPVYKAKYERSIDGQVETITKQGNHEPPONTKGNKONTANING  
SSINNNGSSSELASQFOTVNSMKTREOHEAVSQATTEHLSBADGEGVNGENTDY  
KVEDENPDPRKSTERYISEPAPASHRTVPRATVITVQTSVVDLDDGVRMRKQGO  
KYVGNGNYPSPSYKCTPGCGVARKHVRATIDKRAVITTEGKHNDLPAKSSSHAA  
AAALRDNRNPGGLANNOOQOOPARLRLKEEQTT"  
complement(22794..24934)  
/gene="F7A19.6"  
complement(join(22794..22904,23002..23124,23218..23412,  
23491..23791,23882..24020,24112..24381,24466..24552,  
24646..24934))  
/gene="F7A19.6"  
/note="Hypothetical protein"  
/codon\_start=1  
/protein\_id="AAd39283.1"  
/db\_xref="GI:5080773"  
/translation="MGGCVSQRKLSNKLQOKKOKGRSGCSKISASMPVPMKRM  
SNASVDRFANSEYVHLDPNCAKMKCKRAEMSNANFHLTLOQMNCSQIDGNATISHEE  
AWYDSFYIDSDSDGNSNVFEDANASAGQVIOYEFEESTLKIDGNNGEYSSKN  
EVSIRKQVADSESHHETFTTTCEDHODHRKCKSVVMSVSRSTSIDSKSTPSEFSSA





## KEYWORDS

HTG; HTGS\_PHASE1.  
fruit fly.

## SOURCE

Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.REFERENCE  
AUTHORS1 (bases 1 to 169457)  
Celniker, S.E., Abmayyan, A., Arcalana, T.T., Baxter, E., Blazej, R.G.,  
Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,  
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,  
Hinkle, C.M., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K.,  
Kearney, L., Lee, B., Lewis, S., Li, P., Ling, H., Moshrefi, A.R.,  
Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B.,  
Richards, S., Sethi, H., Svirskas, R.R., Man, K.H., Webster, D.,  
Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G.M.  
Sequencing of Drosophila melanogaster  
Unpublished  
2 (bases 1 to 169457)  
Celniker, S.E., Abmayyan, A., Arcalana, T.T., Baxter, E., Blazej, R.G.,  
Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,  
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,  
Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L.,  
Kim, E., Lee, B., Lewis, S., Li, P., Lomolan, M.A., Mazda, P.,  
Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S.,  
Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Snir, E.,  
Svirskas, R.R., Man, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and  
Rubin, G.M.TITLE  
JOURNAL  
REFERENCE  
AUTHORSDirect Submission  
Submitted (06-ADU-1999) Drosophila Genome Center, Lawrence Berkeley  
Laboratory, MS 64-121, Berkeley, CA 94720, USA  
On Dec 3, 1999 this sequence version replaced gl:5478893.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our sequence  
archive web site (<http://www.fruitfly.org/sequence/>) or send email  
to [bdg@fruitfly.berkeley.edu](mailto:bdg@fruitfly.berkeley.edu). All contigs in this submission meet  
the following cutoffs: length >= 200 bases.

## COMMENT

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 118 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.\* 1  
\* 490 489: contig of 489 bp in length  
\* 570 569: gap of unknown length  
\* 1127 1126: contig of 557 bp in length  
\* 1207 1206: gap of unknown length  
\* 1562 1561: contig of 355 bp in length  
\* 1642 1641: gap of unknown length  
\* 2170 2169: contig of 528 bp in length  
\* 2250 2249: gap of unknown length  
\* 2834 2833: contig of 584 bp in length  
\* 2914 2913: gap of unknown length  
\* 3802 3802: contig of 889 bp in length  
\* 3883 3882: gap of unknown length  
\* 4829 4828: contig of 946 bp in length  
\* 4909 4908: gap of unknown length  
\* 5574 5573: contig of 665 bp in length  
\* 5654 5653: gap of unknown length  
\* 6273 6272: contig of 619 bp in length  
\* 6353 6352: gap of unknown length  
\* 7268 7267: contig of 915 bp in length  
\* 7348 7347: gap of unknown length  
\* 7901 7900: contig of 553 bp in length  
\* 7981 7980: gap of unknown length  
\* 9023 9022: contig of 1042 bp in length  
\* 9103 9102: gap of unknown length  
\* 9693 9692: contig of 590 bp in length  
\* 9773 9772: gap of unknown length  
\* 10395 10394: contig of 622 bp in length  
\* 10475 10474: gap of unknown length  
\* 10877 10877: contig of 403 bp in length\* 10878 10957: gap of unknown length  
\* 10958 11712: contig of 755 bp in length  
\* 11713 11712: gap of unknown length  
\* 11793 12174: contig of 382 bp in length  
\* 12175 12254: gap of unknown length  
\* 12255 12898: contig of 644 bp in length  
\* 12899 12979: gap of unknown length  
\* 12979 14013: contig of 1035 bp in length  
\* 14014 14093: gap of unknown length  
\* 14094 15065: contig of 972 bp in length  
\* 15066 15145: gap of unknown length  
\* 15146 15914: contig of 769 bp in length  
\* 15915 15994: gap of unknown length  
\* 15995 16419: contig of 425 bp in length  
\* 16420 16499: gap of unknown length  
\* 16500 17461: contig of 962 bp in length  
\* 17462 17541: gap of unknown length  
\* 17542 18098: contig of 557 bp in length  
\* 18099 18178: gap of unknown length  
\* 18179 18988: contig of 810 bp in length  
\* 18989 19068: gap of unknown length  
\* 19069 20212: contig of 1144 bp in length  
\* 20213 20292: gap of unknown length  
\* 20293 20817: contig of 525 bp in length  
\* 20818 20897: gap of unknown length  
\* 20898 22197: contig of 1300 bp in length  
\* 22198 22277: gap of unknown length  
\* 22278 23438: contig of 1161 bp in length  
\* 23439 23518: gap of unknown length  
\* 23519 24889: contig of 1371 bp in length  
\* 24890 24970: gap of unknown length  
\* 24970 25810: contig of 841 bp in length  
\* 25811 25890: gap of unknown length  
\* 25891 27149: contig of 1259 bp in length  
\* 27150 27229: gap of unknown length  
\* 27230 28261: contig of 1032 bp in length  
\* 28262 28341: gap of unknown length  
\* 28342 29712: contig of 1371 bp in length  
\* 29713 29792: gap of unknown length  
\* 29793 31385: contig of 1533 bp in length  
\* 31386 31465: gap of unknown length  
\* 31466 32803: contig of 1338 bp in length  
\* 32804 32883: gap of unknown length  
\* 32884 34706: contig of 1823 bp in length  
\* 34707 34786: gap of unknown length  
\* 34787 36436: contig of 1650 bp in length  
\* 36437 36516: gap of unknown length  
\* 36517 38395: contig of 1879 bp in length  
\* 38396 38475: gap of unknown length  
\* 38476 39602: contig of 1127 bp in length  
\* 39603 39682: gap of unknown length  
\* 39683 41589: contig of 1907 bp in length  
\* 41590 41669: gap of unknown length  
\* 41670 43475: contig of 1806 bp in length  
\* 43476 43555: gap of unknown length  
\* 43556 45795: contig of 2240 bp in length  
\* 45796 45875: gap of unknown length  
\* 45876 48574: contig of 2639 bp in length  
\* 48575 48654: gap of unknown length  
\* 48655 50881: contig of 2227 bp in length  
\* 50882 50961: gap of unknown length  
\* 50962 52818: contig of 1857 bp in length  
\* 52819 52898: gap of unknown length  
\* 52899 55256: contig of 2358 bp in length  
\* 55337 55336: gap of unknown length  
\* 55337 59040: contig of 3704 bp in length  
\* 59041 59120: gap of unknown length  
\* 59121 62581: contig of 3461 bp in length  
\* 62582 62661: gap of unknown length  
\* 62662 66265: contig of 3604 bp in length  
\* 66266 71526: gap of unknown length  
\* 71526 71606: gap of unknown length

```

* 71607 76897: contig of 5291 bp in length
* 76898 76977: gap of unknown length
* 76978 83732: contig of 6755 bp in length
* 83733 83812: gap of unknown length
* 83813 89149: contig of 5337 bp in length
* 89150 89229: gap of unknown length
* 89230 97652: contig of 8423 bp in length
* 97653 97732: gap of unknown length
* 97733 106244: contig of 8512 bp in length
* 106245 106325: gap of unknown length
* 106326 116692: contig of 10368 bp in length
* 116693 116772: gap of unknown length
* 116773 129843: contig of 13071 bp in length
* 129844 129923: gap of unknown length
* 129924 130419: contig of 496 bp in length
* 130420 130499: gap of unknown length
* 130500 131145: contig of 646 bp in length
* 131146 131225: gap of unknown length
* 131226 131840: contig of 615 bp in length
* 131841 131920: gap of unknown length
* 131921 132549: contig of 629 bp in length
* 132550 132629: gap of unknown length
* 132630 133129: contig of 500 bp in length
* 133130 133209: gap of unknown length
* 133210 133458: contig of 249 bp in length
* 133459 133538: gap of unknown length
* 133539 134252: contig of 714 bp in length
* 134253 134332: gap of unknown length
* 134333 134855: contig of 523 bp in length
* 134856 134935: gap of unknown length
* 134936 135396: contig of 461 bp in length
* 135397 135476: gap of unknown length
* 135477 135988: contig of 512 bp in length
* 135989 136068: gap of unknown length
* 136069 136855: contig of 787 bp in length
* 136856 136935: gap of unknown length
* 136936 137763: contig of 828 bp in length
* 137764 137843: gap of unknown length
* 137844 138304: contig of 461 bp in length
* 138305 138385: gap of unknown length
* 138386 138905: contig of 521 bp in length
* 138906 138985: gap of unknown length
* 138986 139626: contig of 641 bp in length
* 139627 139706: gap of unknown length
* 139707 140425: contig of 719 bp in length
* 140426 140505: gap of unknown length
* 140506 141245: contig of 740 bp in length
* 141246 141325: gap of unknown length

```

Query Match 11.28; Score 116; DB 61; Length 169457;  
 Best Local Similarity 54.48; Pred. No. 9.3e-22;  
 Matches 233; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

```

OY 415 ccccttggaaatcaatgcatgagggatgctgctgtattgacacattttacc 474
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122413 CCGGTTTGGAAGTGGCCCTGTGGGCTGACGGCGGAGCCGTTGCCACTGCTTGC 122472
OY 475 aatcaactgactagttcaaggttcaagatgaatgaaggaagaaagaaactggaaga 534
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122473 TCGCCGCTGACGTGCAAGTGAAGTGAATCAANTGAGGCGGCGACGCTGATGGGC 122532
OY 535 aaacattggaattcgtgtgtacatcatgacatttgcacaaatcctagctgaagagga 594
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122533 GAGCCACCGAGGTCGACATCCGCTGGCCATGCTTCCGCCAGATCGTCACGTCGGCGGA 122592
OY 595 ataagagaggtcttggcagaggtgtgtgtacccaatatataaagaagcagactgtgtaattg 654
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122593 ATTAAAGGCGCTGTGGAAAGGAGGATCCCGAATGTCAGCGGCGCTTGTGCAATCTG 122652
OY 655 ggaagatttaacacattatgatacagtgaaacacactctgtgattgaaatacaacacttgag 714
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122653 GGGGACCTTAACACATAGACATCAAGCACTGATCATGATGATGCGCTGCAGATGCCC 122712

```

```

OY 715 gacaaatcatgactacaggtttatcaagttatgttcgtgactgtagcttctatctc 774
|| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122713 GACTGTACACAGTACAGTCTGCTGCGCTGCGCTGCTTGTGGCGGATTCGTCAGCATCATG 122772
OY 775 ggaacaccagccgcatgtcatcaaaagcgaatataatgataccacagagataaacaaga 834
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122773 GGCACGCCAGCTGATGTGTGAAGACGCGCATCATGAACGCCACCGACGAGAAATGGC 122832
OY 835 agggagact 842
|| || |||||
Db 122833 CGGTAAGT 122840

RESULT 13
AC012162/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (21-OCT-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Feb 11, 2000 this sequence version replaced q1:6838825.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 142 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
457 456: contig of 456 bp in length
537 536: gap of unknown length
1005 1084: contig of 468 bp in length
1085 1084: gap of unknown length
1959 1958: contig of 874 bp in length
2039 2038: gap of unknown length
2248 2248: contig of 210 bp in length
2249 2248: gap of unknown length

```

2329 2858: contig of 530 bp in length  
2859 2938: gap of unknown length  
2939 3325: contig of 387 bp in length  
3326 3405: gap of unknown length  
3406 4165: contig of 760 bp in length  
4166 4245: gap of unknown length  
4246 4716: contig of 471 bp in length  
4717 4796: gap of unknown length  
4797 5373: contig of 577 bp in length  
5374 5453: gap of unknown length  
5454 6050: contig of 597 bp in length  
6051 6130: gap of unknown length  
6131 6720: contig of 590 bp in length  
6721 6800: gap of unknown length  
6801 7461: contig of 661 bp in length  
7462 7541: gap of unknown length  
7542 8104: contig of 563 bp in length  
8105 8184: gap of unknown length  
8185 8505: contig of 321 bp in length  
8506 8585: gap of unknown length  
8586 9227: contig of 642 bp in length  
9228 9307: gap of unknown length  
9308 10091: contig of 784 bp in length  
10092 10171: gap of unknown length  
10172 11099: contig of 928 bp in length  
11100 11179: gap of unknown length  
11180 11747: contig of 568 bp in length  
11748 11827: gap of unknown length  
11828 12767: contig of 940 bp in length  
12768 12847: gap of unknown length  
12848 13398: contig of 551 bp in length  
13399 13478: gap of unknown length  
13479 14007: contig of 529 bp in length  
14008 14087: gap of unknown length  
14088 14867: contig of 780 bp in length  
14868 14947: gap of unknown length  
14948 15562: contig of 615 bp in length  
15563 15642: gap of unknown length  
15643 16859: contig of 1217 bp in length  
16860 16939: gap of unknown length  
16940 17986: contig of 1047 bp in length  
17987 18066: gap of unknown length  
18067 19341: contig of 1275 bp in length  
19342 19421: gap of unknown length  
19422 20926: contig of 1505 bp in length  
20927 21006: gap of unknown length  
21007 22153: contig of 1147 bp in length  
22154 22233: gap of unknown length  
22234 23236: contig of 1003 bp in length  
23237 23316: gap of unknown length  
23317 24607: contig of 1291 bp in length  
24608 24687: gap of unknown length  
24688 25620: contig of 933 bp in length  
25621 25700: gap of unknown length  
25701 26300: contig of 600 bp in length  
26301 26380: gap of unknown length  
26381 27488: contig of 1108 bp in length  
27489 27568: gap of unknown length  
27569 28215: contig of 647 bp in length  
28216 28295: gap of unknown length  
28296 29440: contig of 1145 bp in length  
29441 29520: gap of unknown length  
29521 30587: contig of 1067 bp in length  
30588 30667: gap of unknown length  
30668 32042: contig of 1375 bp in length  
32043 32122: gap of unknown length  
32123 33740: contig of 1618 bp in length  
33741 33820: gap of unknown length  
33821 35375: contig of 1555 bp in length  
35376 35455: gap of unknown length  
35456 36258: contig of 803 bp in length  
36259 36338: gap of unknown length  
36339 38003: contig of 1665 bp in length

38004 38083: gap of unknown length  
38084 39509: contig of 1426 bp in length  
39510 39589: gap of unknown length  
39590 40719: contig of 1130 bp in length  
40720 40799: gap of unknown length  
40800 42453: contig of 1654 bp in length  
42454 42533: gap of unknown length  
42534 43841: contig of 1308 bp in length  
43842 43921: gap of unknown length  
43922 45410: contig of 1489 bp in length  
45411 45490: gap of unknown length  
45491 46782: contig of 1292 bp in length  
46783 46862: gap of unknown length  
46863 48570: contig of 1708 bp in length  
48571 48650: gap of unknown length  
48651 51042: contig of 2392 bp in length  
51043 51122: gap of unknown length  
51123 52760: contig of 1638 bp in length  
52761 52840: gap of unknown length  
52841 54712: contig of 1872 bp in length  
54713 54792: gap of unknown length  
54793 56521: contig of 1729 bp in length  
56522 56601: gap of unknown length  
56602 58938: contig of 2337 bp in length  
58939 59018: gap of unknown length  
59019 60837: contig of 1819 bp in length  
60838 60917: gap of unknown length  
60918 62968: contig of 2071 bp in length  
62969 63068: gap of unknown length  
63069 65583: contig of 2515 bp in length  
65584 65663: gap of unknown length  
65664 68774: contig of 3111 bp in length  
68775 68854: gap of unknown length  
68855 71235: contig of 2381 bp in length  
71236 71315: gap of unknown length  
71316 74193: contig of 2798 bp in length  
74194 74194: gap of unknown length  
74195 77405: contig of 3212 bp in length  
77406 77485: gap of unknown length  
77486 80936: contig of 3451 bp in length  
80937 81016: gap of unknown length  
81017 83357: contig of 2341 bp in length  
83358 84437: gap of unknown length  
84438 87456: contig of 4019 bp in length  
87457 87536: gap of unknown length  
87537 91532: contig of 3996 bp in length  
91533 91612: gap of unknown length  
91613 97143: contig of 5531 bp in length  
97144 97223: gap of unknown length  
97224 105593: contig of 8370 bp in length  
105594 105673: gap of unknown length  
105674 112071: contig of 6398 bp in length  
112071 112071: gap of unknown length  
112072 121091: contig of 8860 bp in length  
121091 121091: gap of unknown length  
121092 143113: contig of 22022 bp in length  
143113 143193: gap of unknown length  
143193 143587: contig of 394 bp in length  
143587 143667: gap of unknown length  
143668 143927: contig of 260 bp in length  
143927 144007: gap of unknown length  
144007 144842: contig of 835 bp in length  
144842 144922: gap of unknown length  
144922 145360: contig of 438 bp in length  
145360 145440: gap of unknown length  
145440 145461: contig of 475 bp in length  
145461 145916: gap of unknown length  
145916 145996: contig of 537 bp in length  
145996 146532: gap of unknown length  
146532 146533: gap of unknown length

Query Match 11.2%; Score 116; DB 63; Length 191504;  
Best Local Similarity 54.4%; Pred. No. 9; Ae-22;  
Matches 233; Conservative 0; Mismatches 195; Indels 0; Gaps 0;















GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2001, 17:32:31 ; Search time 112.35 seconds  
(without alignments)  
5398.743 Million cell updates/sec

Title: US-09-397-342-2

Perfect score: 1039  
Sequence: 1 ccgagctccgagctccgcttat.....cagatatccatccactgc 1039

Scoring table: IDENTITY\_NUC  
Gap 10.0 , Gapext 1.0

Searched: 678276 seqs, 291890651 residues

Total number of hits satisfying chosen parameters: 1356552

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N.GeneSeq\_0401.\*  
1: /SID6/gcgdata/geneSeq/geneSeq/NA1980.DAT.\*  
2: /SID6/gcgdata/geneSeq/geneSeq/NA1981.DAT.\*  
3: /SID6/gcgdata/geneSeq/geneSeq/NA1982.DAT.\*  
4: /SID6/gcgdata/geneSeq/geneSeq/NA1983.DAT.\*  
5: /SID6/gcgdata/geneSeq/geneSeq/NA1984.DAT.\*  
6: /SID6/gcgdata/geneSeq/geneSeq/NA1985.DAT.\*  
7: /SID6/gcgdata/geneSeq/geneSeq/NA1986.DAT.\*  
8: /SID6/gcgdata/geneSeq/geneSeq/NA1987.DAT.\*  
9: /SID6/gcgdata/geneSeq/geneSeq/NA1988.DAT.\*  
10: /SID6/gcgdata/geneSeq/geneSeq/NA1989.DAT.\*  
11: /SID6/gcgdata/geneSeq/geneSeq/NA1990.DAT.\*  
12: /SID6/gcgdata/geneSeq/geneSeq/NA1991.DAT.\*  
13: /SID6/gcgdata/geneSeq/geneSeq/NA1992.DAT.\*  
14: /SID6/gcgdata/geneSeq/geneSeq/NA1993.DAT.\*  
15: /SID6/gcgdata/geneSeq/geneSeq/NA1994.DAT.\*  
16: /SID6/gcgdata/geneSeq/geneSeq/NA1995.DAT.\*  
17: /SID6/gcgdata/geneSeq/geneSeq/NA1996.DAT.\*  
18: /SID6/gcgdata/geneSeq/geneSeq/NA1997.DAT.\*  
19: /SID6/gcgdata/geneSeq/geneSeq/NA1998.DAT.\*  
20: /SID6/gcgdata/geneSeq/geneSeq/NA1999.DAT.\*  
21: /SID6/gcgdata/geneSeq/geneSeq/NA2000.DAT.\*  
22: /SID6/gcgdata/geneSeq/geneSeq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1039	100.0	1039	21	Al4084
2	998	96.1	998	21	A37139
3	990.6	95.3	1248	21	Human uncoupling p
4	676.8	65.1	1726	22	F27721
5	286.8	27.6	828	21	C76281
6	140	13.5	1069	21	C46432
7	140	13.5	1071	21	C39993
8	124	11.9	1240	21	C45636
9	124	11.9	1242	21	C40226
10	113.4	10.9	1103	21	C46816
11	106	10.2	1227	21	C33096

12	104.6	10.1	1224	21	C49402	Arabidopsis thalia
13	97.2	9.4	957	21	C49389	Arabidopsis thalia
14	93.4	9.0	876	21	A37719	Novel human protei
15	89.6	8.6	1031	21	A27999	Murine uncoupling
16	88	8.5	1022	21	A27998	Murine uncoupling
17	79.6	7.7	936	21	A27995	Human uncoupling p
18	78.4	7.5	993	21	A27992	Human uncoupling p
19	78.4	7.5	1009	21	A27996	Human uncoupling p
20	78.4	7.5	1674	21	A29803	Human secreted pro
21	77.8	7.5	1178	21	C49924	Arabidopsis thalia
22	77.8	7.5	1180	21	C42453	Arabidopsis thalia
23	76.8	7.4	978	21	A27995	Human uncoupling p
24	76.2	7.3	1416	21	C38589	Arabidopsis thalia
25	74.2	7.1	1411	21	C50232	Arabidopsis thalia
26	63	6.1	1166	21	C34840	Arabidopsis thalia
27	60.8	5.9	1398	20	X25083	Potato tuber-speci
28	57.2	5.5	924	21	A29243	Human uncoupling p
29	57.2	5.5	924	21	A29243	Human uncoupling p
30	56.2	5.4	882	21	A37720	Novel human protei
31	52.6	5.1	1103	21	A27997	Human uncoupling p
32	52.6	5.1	1430	21	A26374	Human secreted pro
33	52.4	5.0	1204	19	V71712	Mouse uncoupling p
34	51.8	5.0	960	22	C83793	Hybrid hUCP2 DNA.
35	50.8	4.9	1658	20	V84307	Mouse uncoupling p
36	50.8	4.9	1949	20	Z07061	Mouse uncoupling p
37	50.8	4.9	1949	20	V71228	CDNA encoding a mu
38	50.8	4.9	1949	21	Z90319	CDNA encoding a mu
39	49.4	4.8	930	20	Z19968	Human uncoupling p
40	49.4	4.8	930	20	X99435	Human uncoupling p
41	49.4	4.8	930	21	Z50624	Human uncoupling p
42	49.4	4.8	1105	20	X99434	UCP2 Nucleotide se
43	49.4	4.8	1105	21	Z29323	Human uncoupling p
44	49.4	4.8	1255	17	T13981	Human body weight
45	49.4	4.8	1255	19	V09078	Human C5 gene CDNA

## ALIGNMENTS

RESULT	1
Al4084	Al4084 standard; CDNA: 1039 BP.
XX	
AC	Al4084;
XX	
DT	21-JUL-2000 (first entry)
XX	
DE	CDNA encoding human uncoupling protein UCP4.
XX	
KW	Uncoupling protein; UCP4; expressed sequence tag; EST; human;
KW	mitochondrial membrane; ATP synthesis; energy efficiency;
KW	drug screening; obesity; stroke; trauma; sepsis; infection; ss.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	Location/Qualifiers
FT	CDS
XX	40..1011
XX	/*tag= a
XX	/product= "Human UCP4"
PN	MO200017353-A1.
XX	
PD	30-MAR-2000.
XX	
PF	15-SEP-1999; 99WO-US21194.
XX	
PR	22-SEP-1998; 98US-0101279.
PR	30-DEC-1998; 98US-0114223.
PR	16-APR-1999; 99US-0129674.
XX	
XX	
PA	(GETH ) GENENTECH INC.



```
PR 09-SEP-1998; 98US-0099596.
PR 09-SEP-1998; 98US-0099598.
PR 09-SEP-1998; 98US-0099602.
PR 09-SEP-1998; 98US-0099642.
PR 10-SEP-1998; 98US-0099741.
PR 10-SEP-1998; 98US-0099754.
PR 10-SEP-1998; 98US-0099763.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.
PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 10-SEP-1998; 98US-0099816.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 15-SEP-1998; 98US-0100394.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.

PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.
PR 26-OCT-1998; 98US-0105694.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105811.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0108500.
PR 30-OCT-1998; 98US-0108464.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106902.
PR 03-NOV-1998; 98US-0106905.
PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108779.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 18-NOV-1998; 98US-0108925.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108849.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108851.
PR 18-NOV-1998; 98US-0108852.
PR 18-NOV-1998; 98US-0108858.
PR 18-NOV-1998; 98US-0108904.

PA (GETH ) GENENTECH INC.
XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WT;
PI
XX WPI: 2000-237871/20.
XX P-PSDB; Y99457.
XX
XX New mammalian DNA sequences encoding transmembrane, receptor or
PT secreted PRO polypeptides, useful for screening of potential peptide or
PT small molecule inhibitors of the relevant receptor/ligand interactions
XX
XX Claim 2; Fig 235; 773pp; English.
XX
XX A37022 to A37144 encode the new isolated human transmembrane, receptor
CC or secreted PRO polypeptides given in Y99340 to Y99462. The
CC transmembrane and receptor PRO proteins can be used for screening of
CC potential peptide or small molecule inhibitors of the relevant
CC receptor/ligand interactions. The polypeptides and nucleotide sequences
CC encoding them have various industrial applications, including uses as
CC pharmaceutical and diagnostic agents. A37145 to A37330 represent
CC PCR primers and hybridisation probes used in the isolation of the PRO
CC polypeptides from the present invention.
XX
XX Sequence 998 BP; 284 A; 202 C; 262 G; 250 T; 0 other;
SO
```

Query Match 96.1%; Score 998; DB 21; Length 998;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 998; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 ccgtatcgtcttgcgtactctgtaatgcccgtccggagagagagaggtttg 73  
 DB 1 ccgtatcgtcttgcgtactctgtaatgcccgtccggagagagagaggtttg 60  
 QY 74 ccgtacccagagatgcccggagagagagagagagagagagagagagagag 133  
 DB 61 ccgtacccag 120  
 QY 134 tggcagagatgagacaccttcccttgatctcaaaaactgactcccaatgcaagag 193  
 DB 121 tggcagagatgagacaccttcccttgatctcaaaaactgactcccaatgcaagag 180  
 QY 194 aagcagctctgctggttggagacggtgcaagaaatctgcccctaagggagatg 253  
 DB 181 aagcagctctgctggttggagacggtgcaagaaatctgcccctaagggagatg 240  
 QY 254 tggcagacagccctgaagatctggaagagagagagagagagagagagagagatg 313  
 DB 241 tggcagacagccctgaagatctggaagagagagagagagagagagagagagatg 300  
 QY 314 caaccgcaattacagacagcagtagtattctgagagtcgaaatggtacataatgacatc 373  
 DB 301 caaccgcaattacagacagcagtagtattctgagagtcgaaatggtacataatgacatc 360  
 QY 374 tccgagaggtgtgtttggcaaaagtgaatgagcatatccctcttggaaatcagta 433  
 DB 361 tccgagaggtgtgtttggcaaaagtgaatgagcatatccctcttggaaatcagta 420  
 QY 434 tggagagagatggtgtgtttggcaaaagtgaatgagcatatccctcttggaaatcagta 493  
 DB 421 tggagagagatggtgtgtttggcaaaagtgaatgagcatatccctcttggaaatcagta 480  
 QY 494 aggttcaagatgcaaaatggaagagagagagagagagagagagagagagagagagagag 553  
 DB 481 aggttcaagatgcaaaatggaagagagagagagagagagagagagagagagagagagag 540  
 QY 554 ggtacatcatgcatcttgcgaagagagagagagagagagagagagagagagagagagag 613  
 DB 541 ggtacatcatgcatcttgcgaagagagagagagagagagagagagagagagagagagag 600  
 QY 614 gctgggtaccacatatacaagagcagcctgtgaatgagagagagagagagagagagagagag 673  
 DB 601 gctgggtaccacatatacaagagcagcctgtgaatgagagagagagagagagagagagagag 660  
 QY 674 atacagtgaaacactactgtgatgaatatacaccacttgagagacatcatgactcag 733  
 DB 661 atacagtgaaacactactgtgatgaatatacaccacttgagagacatcatgactcag 720  
 QY 734 gtttcaagttatgttctgagcgtgtgcttctatcttggagacacagccagatgta 793  
 DB 721 gtttcaagttatgttctgagcgtgtgcttctatcttggagacacagccagatgta 780  
 QY 794 tcaaaagcagaataatgaaatcaaccagagatacaaaagagagagagagagagagagagagagag 853  
 DB 781 tcaaaagcagaataatgaaatcaaccagagatacaaaagagagagagagagagagagagagagag 840  
 QY 854 catcagactgactgtgactcagcgtgtcaaggtgaagatcatgactatataaag 913  
 DB 841 catcagactgactgtgactcagcgtgtcaaggtgaagatcatgactatataaag 900  
 QY 914 gtttttaccatcttgcgtgag 973  
 DB 901 gtttttaccatcttgcgtgag 960  
 QY 974 aaaaaatcag 1011  
 DB 961 aaaaaatcag 998  
 RESULT 3  
 ID A14085  
 XX A14085 standard; DNA: 1248 BP.

AC A14085;  
 XX 21-JUL-2000 (first entry)  
 DE Human uncoupling protein UCP4 "from DNA" sequence, SEQ ID NO:5.  
 XX  
 KW Uncoupling protein; UCP4; expressed sequence tag; EST; UCP3 homologue;  
 KW human; chromosome 6p11.2-q12; ATP synthesis; energy efficiency;  
 KW mitochondrial membrane; proton leakage; heat production; metabolic rate;  
 KW drug screening; obesity; stroke; trauma; sepsis; infection; ds.  
 OS Homo sapiens.  
 PM WO200017353-A1.  
 PD 30-MAR-2000.  
 XX 15-SEP-1999; 99WO-US21194.  
 PE 22-SEP-1998; 98US-0101279.  
 PR 30-DEC-1998; 98US-0114223.  
 PR 16-APR-1999; 99US-0129674.  
 XX  
 PA (GERTH ) GENENTECH INC.  
 XX  
 PI Adams S, Pan J, Zhong A;  
 XX  
 DR WPI: 2000-292842/25.  
 XX  
 PT New nucleic acid encoding human uncoupled protein-4, useful e.g for  
 PT identifying metabolic regulators for treatment of obesity -  
 XX  
 PS Example 1; Fig 7; 80pp; English.  
 XX  
 CC The invention relates to human uncoupling protein UCP4 (Y81497) and  
 CC cDNA encoding it (A14085). The human UCP4 cDNA (Accession 200334) was isolated  
 CC from a brain cDNA library using a probe generated using primers  
 CC A14086-A14087. These primers were based on a UCP4 "from DNA" sequence  
 CC (A14085) derived from a number of ESTs (expressed sequence tags) which  
 CC were selected on the basis of homology with human UCP3. The human UCP4  
 CC gene has been mapped to chromosome 6p11.2-q12. UCP4 catalyses the leakage  
 CC of protons through the mitochondrial membrane, thus bypassing ATP  
 CC synthase and thereby reducing the efficiency of ATP synthesis. Modulation  
 CC of UCP4 activity or expression can therefore alter the metabolic rate and  
 CC heat production via modulation of ATP synthetic efficiency. UCP4 nucleic  
 CC acids may be used for recombinant production of UCP4 and as a source of  
 CC primers and hybridisation probes which may be used for the analysis of  
 CC UCP4 expression, for screening for homologous sequences, and for  
 CC chromosome or gene mapping. They can be also be used to produce  
 CC transgenic or knockout animals for the development and screening of  
 CC therapeutic agents, as a source of antisense nucleotides, and in gene  
 CC therapy for metabolic disorders. The UCP4 protein can be used to screen  
 CC for specific modulators of activity and to raise antibodies which may be  
 CC used therapeutically as UCP4 antagonists or UCP4 targeting proteins, for  
 CC affinity purification of UCP4, and as immunoassay reagents for detecting  
 CC UCP4 expression, e.g., for the diagnosis of impaired neural activity or  
 CC neural degeneration. Agents that modulate UCP4 activity are used to  
 CC control the metabolic rate in mammals. UCP4 upregulators can be used to  
 CC increase the metabolic rate in order to treat obesity and the symptoms  
 CC associated with stroke, trauma, sepsis and infection. The present  
 CC sequence represents the human UCP4 "from DNA" sequence assembled from EST  
 CC sequences with homology to human UCP3.  
 XX  
 SQ Sequence 1248 BP; 357 A; 263 C; 301 G; 326 T; 1 other:  
 QY Query Match 95.3%; Score 990.6; DB 21; Length 1248;  
 QY Best Local Similarity 99.6%; Pred. No. 0;  
 QY Matches 993; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 DB 15 ccgtatcgtcttgcgtactctgtaatgcccgtccggagagagagagagagagagagagagagag 74  
 DB 1 ccgtatcgtcttgcgtactctgtaatgcccgtccggagagagagagagagagagagagagagag 60





DB 812 tggatgtgaataacaccacttgaggaacataatcatgactcaggtttatcatatgtt 871  
QY 752 ctggactggtagcttctatctctgggaacacccagctgltcatcaaaagcagataatga 811  
DB 872 ctggactggtagcttctatctctgggaacacccagctgltcatcaaaagcagataatga 931  
QY 812 atcaacacagagataaacaaggaaggagcttctgtataatcatcgaactgactgttga 871  
DB 932 atcaacacagagataaacaaggaaggagcttctgtataatcatcgaactgactgttga 991  
QY 872 ttcaggcttcaaggctgaagattcatgactataaaggctttttaccactgtgc 931  
DB 992 ttcaggcttcaaggctgaagattcatgactataaaggctttttaccactgtgc 1051  
QY 932 tgaagaagacccttggatcaatggtgttctggttactatgaaaaaatacagagagatga 991  
DB 1052 tgaagaagacccttggatcaatggtgttctggttactatgaaaaaatacagagagatga 1111  
QY 992 gtggagtcagtcattttaa 1011  
1112 gtggagtcagtcattttaa 1131

## RESULT 5

C76281 ID C76281 standard; cDNA: 828 BP.

AC C76281;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF1836 polynucleotide sequence SEQ ID NO:3671.

XX Human; open reading frame; ORFX: detection; cytosolic; hepatotropic;  
XX vulerary; antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;  
XX anticonvulsant; osteoplastic; antidiabetic; immunosuppressant; cardiant;  
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
XX hypotensive; dermatological; immunosuppressive; antineoplastic;  
XX antiviral; antibacterial; antifungal; antineuritic; antithyroid;  
XX antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;  
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
XX cholesterol ester storage; systemic lupus erythematosus; infection;  
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
XX bone damage; cartilage damage; antineoplastic disease; coagulation;  
XX thrombosis; contraceptive; ss.

KW Homo sapiens.

PN WO200058473-A2.

PD 05-OCT-2000.

PF 31-MAR-2000; 2000MO-US08621.

XX 31-MAR-1999; 990US-0127607.

PR 02-APR-1999; 990US-0127636.

PR 05-APR-1999; 990US-0127728.

PR 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shinkels RA, Leach M;

DR WPI: 2000-602362/57.

XX P-PSDB; BA2072.

XX Novel nucleic acids and peptides derived from open reading frame X,  
XX useful for treating e.g. cancers, proliferative disorders,  
XX neurodegenerative disorders and cardiovascular disease -  
PS Claim 5; Page 2828; 5507pp; English.

XX C74446 to C77606 encode the proteins given in B40237 to B43397, which  
CC represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytosolic; hepatotropic; vulerary;  
CC antiproliferative; antiparkinsonian; neurotrophic; neuroprotective; osteoplastic;  
CC anticonvulsant; antidiabetic; immunosuppressant; immunostimulant;  
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
CC dermatological; immunosuppressive; antineoplastic; antithyroid;  
CC antiviral; antifungal; antineuritic; antidiabetic; antianemic. The  
CC sequences can be used for determining the presence of or predisposition  
CC to, or preventing or treating pathological conditions associated with an  
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
CC used to treat cancers, proliferative disorders, neurodegenerative  
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
CC cartilage damage, nocturnal haemoglobinuria, antineoplastic disease; to  
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive.  
XX

SQ Sequence 828 BP; 276 A; 133 C; 201 G; 216 T; 2 other:

Query Match 27.6%; Score 286.8; DB 21; Length 828;  
Best Local Similarity 99.3%; Pred. No. 8.1e-84;  
Matches 288; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 332 acgtatgtatctctgaggtggaatggtacatgaacatccgagaggtgtgtgt 391  
DB 539 accaagtgatctcggaggtcgaatggtacatgaacatccgagaggtgtgtgt 598  
QY 392 gcaaaagtgaagatgagcatlacccttggaaatcaatcatctgagagatgagctg 451  
DB 599 gcaaaagtgaagatgagcatlacccttggaaatcaatcatctgagagatgagctg 658  
QY 452 gtgtatgtgcaagtttttagccaatccaactgacctagtgaggttccagatgcaatg 511  
DB 659 gtgtatgtgcaagtttttagccaatccaactgacctagtgaggttccagatgcaatg 718  
QY 512 aaggaaaaggaagaaactggaaggaagaaccatctgcatctgtgtgttaccatgtcatgt 571  
DB 719 aaggaaaaggaagaaactggaaggaagaaccatctgcatctgtgtgttaccatgtcatgt 778  
QY 572 caaaatcttaagctgaaggaagaatcagaggtcttggcagcggtgta 621  
DB 779 caaaatcttaagctgaaggaagaatcagaggtcttggcagcggtgta 828

## RESULT 6

C46432 ID C46432 standard; DNA: 1069 BP.

AC C46432;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 50118.

XX Hybridisation assay; genetic mapping; gene expression control;

XX protein identification; signal transduction pathway;

XX metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 990US-0121825.

PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126284.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 18-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139753.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 02-JUL-1999; 99US-0142154.  
 PR 06-JUL-1999; 99US-0142055.  
 PR 08-JUL-1999; 99US-0142390.  
 PR 09-JUL-1999; 99US-0142803.  
 PR 12-JUL-1999; 99US-0142920.  
 PR 13-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144332.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 21-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 06-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0150566.  
 PR 25-AUG-1999; 99US-0150567.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.

PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161970.
PR	28-OCT-1999;	99US-0161982.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match	13.5%	Score 140;	DB 21;	Length 1069;
Best Local Similarity	51.8%;	Pred. No. 1.5e-35;		
Matches 382; Conservative	0;	Mismatches 335;	Indels 21;	Gaps 2

OY	279	aaagaaagagcttctttaaagcttcttgcaagagatgaaaccccgcatattcaagacagta	338
Db	224	aaagaaagagagtgatgtgtctctccaaagttctatctccggcgatattcaagacactglt	283
OY	339	gtattctgtagagtgcaatggtlcaacatatgaaactccag-----aggttggttgg	392
Db	284	ctcacagccctatcagaaatctgtatcagagaattttgaaagactacatcgltcagatctga	343
OY	393	caaaagtgaaagatgatgacattatccctcttggaatcaatcattgtagaggtgtgtgctgg	452
	344	aactcaaaatagcgagctcttctctctcgcgcaaaagccctctgcgagagatttctcgg	403
OY	453	tgthattgcccagtttlltlaagccaactccaactgaactgaatgaaggttcaaatgaa	512
Db	404	tgthattgcccagttgtagtgcagctcagctgatttgttcaaaagtgaattgcaagaga	463
OY	513	agagaaagagaaactgtgaagaaacacattcgatttcggtgtgtacatcatcatcttgc	572
Db	464	tgttagattgtgtgaagccaaagccttgaaacccgaagttactcgggacccaatcgaagctttac	533
OY	573	aaaaactctagctgaagagagaaatagcagaggtcttggcagagcttggttaaccatatca	632
Db	524	caaaactctcaatcagaagaaggttaaaaggttatggaagaagtgcttcttccaactlcca	583
OY	633	aaagacgacacttgtagatactggaggttaaacacattatgatcaggtgaanaacactt	692
Db	584	gagagcacttcttagtgaatactggagaacactagcttgctatgacaacgccaacacttgt	643
OY	693	ggtattgaatacaccaacttgaagagaaatatacatgatacctcaggttatacaagtttatgtc	752
Db	644	catcgataagaagattgtctgtagagataaacatttttgcgcacactcttgcttctaatagtc	703
OY	753	tgtagactgtgacttctatcttctggaaacacagccgatagtatcataaagcgaataatgaa	812
Db	704	gggtcttgcttgtagcaagtttggagttgtccagctgtagtggtgtaaagcagagtgatgaa	763

QY	813	tcaccacgcagatataaacagaaggaaggagactttgtatataatcctgcactgactgctgatt	872
Db	764	ccagagggtgaataatgc-----tctgtataagaatcttctaagcactggtttgt	808
QY	873	tcaagctgtgttcaaggtgaagatcatgagttctataataaagccttttaccatcttgct	932
Db	809	gaagacgctttaaagtttaaggaataagactttgtgtgaaaggtttcttccgcacatgagc	868
QY	933	gagatgaccccttggctcaatggtgtcttctgacttactatgataaaaatcagagagatgag	992
Db	869	aagacttggaccctggtgcagacttcgttttctgtctcctatgagaagattagactgtgac	928
QY	993	tggagtcagtcacattta 1010	
Db	929	aggaatctcttctcttcta 946	
RESULT	7		
ID	C39993	standard; DNA; 1071 BP.	
AC	C39993;		
XX	17-OCT-2000	(first entry)	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 26654.		
XX	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway;		
KW	metabolic pathway; promoter; termination sequence; ss.		
XX	Arabidopsis thaliana.		
OS	EPI033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-0301439.		
XX	25-FEB-1999;	9905-0121825.	
PR	05-MAR-1999;	9905-0123180.	
PR	09-MAR-1999;	9905-0123548.	
PR	23-MAR-1999;	9905-0125788.	
PR	25-MAR-1999;	9905-0126284.	
PR	29-MAR-1999;	9905-0126785.	
PR	01-APR-1999;	9905-0127462.	
PR	06-APR-1999;	9905-0128234.	
PR	08-APR-1999;	9905-0128714.	
PR	16-APR-1999;	9905-0129845.	
PR	19-APR-1999;	9905-0130077.	
PR	21-APR-1999;	9905-0130449.	
PR	23-APR-1999;	9905-0130510.	
PR	23-APR-1999;	9905-0130891.	
PR	28-APR-1999;	9905-0131449.	
PR	30-APR-1999;	9905-0132048.	
PR	30-APR-1999;	9905-0132407.	
PR	04-MAY-1999;	9905-0132484.	
PR	05-MAY-1999;	9905-0132485.	
PR	06-MAY-1999;	9905-0132486.	
PR	06-MAY-1999;	9905-0132487.	
PR	07-MAY-1999;	9905-0132863.	
PR	11-MAY-1999;	9905-0134256.	
PR	14-MAY-1999;	9905-0134218.	
PR	14-MAY-1999;	9905-0134221.	
PR	14-MAY-1999;	9905-0134370.	
PR	18-MAY-1999;	9905-0134768.	
PR	19-MAY-1999;	9905-0134941.	
PR	20-MAY-1999;	9905-0135124.	
PR	21-MAY-1999;	9905-0135353.	
PR	24-MAY-1999;	9905-0135629.	
PR	25-MAY-1999;	9905-0136021.	

PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139452.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139753.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139889.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 30-AUG-1999; 99US-0151080.  
 PR 31-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 26-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158359.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.



PR 30-JUN-1999; 9905-014187.  
PR 01-JUL-1999; 9905-0141842.  
PR 01-JUL-1999; 9905-0142154.  
PR 02-JUL-1999; 9905-0142055.  
PR 06-JUL-1999; 9905-0142390.  
PR 08-JUL-1999; 9905-0142803.  
PR 09-JUL-1999; 9905-0142920.  
PR 12-JUL-1999; 9905-0142977.  
PR 13-JUL-1999; 9905-0143542.  
PR 14-JUL-1999; 9905-0143624.  
PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144333.  
PR 19-JUL-1999; 9905-0144334.  
PR 20-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144632.  
PR 21-JUL-1999; 9905-0144864.  
PR 21-JUL-1999; 9905-0144814.  
PR 21-JUL-1999; 9905-0145086.  
PR 22-JUL-1999; 9905-0145088.  
PR 22-JUL-1999; 9905-0145085.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 22-JUL-1999; 9905-0145192.  
PR 23-JUL-1999; 9905-0145145.  
PR 23-JUL-1999; 9905-0145218.  
PR 23-JUL-1999; 9905-0145224.  
PR 26-JUL-1999; 9905-0145276.  
PR 27-JUL-1999; 9905-0145913.  
PR 27-JUL-1999; 9905-0145918.  
PR 27-JUL-1999; 9905-0145919.  
PR 28-JUL-1999; 9905-0145919.  
PR 02-AUG-1999; 9905-0146386.  
PR 02-AUG-1999; 9905-0146388.  
PR 02-AUG-1999; 9905-0146389.  
PR 03-AUG-1999; 9905-0146389.  
PR 03-AUG-1999; 9905-0147038.  
PR 04-AUG-1999; 9905-0147204.  
PR 04-AUG-1999; 9905-0147302.  
PR 05-AUG-1999; 9905-0147192.  
PR 05-AUG-1999; 9905-0147260.  
PR 06-AUG-1999; 9905-0147303.  
PR 06-AUG-1999; 9905-0147416.  
PR 09-AUG-1999; 9905-0147493.  
PR 09-AUG-1999; 9905-0147935.  
PR 10-AUG-1999; 9905-0148171.  
PR 11-AUG-1999; 9905-0148319.  
PR 12-AUG-1999; 9905-0148341.  
PR 13-AUG-1999; 9905-0148565.  
PR 13-AUG-1999; 9905-0148684.  
PR 16-AUG-1999; 9905-0149368.  
PR 17-AUG-1999; 9905-0149375.  
PR 18-AUG-1999; 9905-0149426.  
PR 20-AUG-1999; 9905-0149723.  
PR 20-AUG-1999; 9905-0149722.  
PR 20-AUG-1999; 9905-0149929.  
PR 23-AUG-1999; 9905-0149902.  
PR 23-AUG-1999; 9905-0149930.  
PR 25-AUG-1999; 9905-0150566.  
PR 26-AUG-1999; 9905-0150884.  
PR 27-AUG-1999; 9905-0151065.  
PR 27-AUG-1999; 9905-0151066.  
PR 27-AUG-1999; 9905-0151080.  
PR 30-AUG-1999; 9905-0151303.  
PR 31-AUG-1999; 9905-0151438.  
PR 01-SEP-1999; 9905-0151930.  
PR 07-SEP-1999; 9905-0152363.  
PR 10-SEP-1999; 9905-0153070.  
PR 13-SEP-1999; 9905-0153758.

PR 15-SEP-1999; 9905-0154018.  
PR 16-SEP-1999; 9905-0154039.  
PR 20-SEP-1999; 9905-0154779.  
PR 22-SEP-1999; 9905-0155139.  
PR 23-SEP-1999; 9905-0155486.  
PR 24-SEP-1999; 9905-0155659.  
PR 28-SEP-1999; 9905-0156458.  
PR 29-SEP-1999; 9905-0156566.  
PR 04-OCT-1999; 9905-0157117.  
PR 05-OCT-1999; 9905-0157173.  
PR 06-OCT-1999; 9905-0157865.  
PR 07-OCT-1999; 9905-0158029.  
PR 08-OCT-1999; 9905-0158232.  
PR 12-OCT-1999; 9905-0158369.  
PR 13-OCT-1999; 9905-0159293.  
PR 13-OCT-1999; 9905-0159294.  
PR 13-OCT-1999; 9905-0159295.  
PR 14-OCT-1999; 9905-0159329.  
PR 14-OCT-1999; 9905-0159330.  
PR 14-OCT-1999; 9905-0159331.  
PR 14-OCT-1999; 9905-0159637.  
PR 14-OCT-1999; 9905-0159638.  
PR 18-OCT-1999; 9905-0159584.  
PR 21-OCT-1999; 9905-0160741.  
PR 21-OCT-1999; 9905-0160767.  
PR 21-OCT-1999; 9905-0160768.  
PR 21-OCT-1999; 9905-0160770.  
PR 21-OCT-1999; 9905-0160814.  
PR 21-OCT-1999; 9905-0160815.  
PR 22-OCT-1999; 9905-0160980.  
PR 22-OCT-1999; 9905-0160981.  
PR 22-OCT-1999; 9905-0160989.  
PR 25-OCT-1999; 9905-0161404.  
PR 25-OCT-1999; 9905-0161405.  
PR 25-OCT-1999; 9905-0161406.  
PR 26-OCT-1999; 9905-0161359.  
PR 26-OCT-1999; 9905-0161360.  
PR 26-OCT-1999; 9905-0161361.  
PR 28-OCT-1999; 9905-0161920.  
PR 28-OCT-1999; 9905-0161992.  
PR 28-OCT-1999; 9905-0161993.  
PR 29-OCT-1999; 9905-0162142.

Query Match 11.98; Score 124; DB 21; Length 1240;  
Best Local Similarity 53.58; Pred. No. 2.9e-30;  
Matches 293; Conservative 0; Mismatches 240; Indels 15; Gaps 1;  
QY 463 cagttttagccaactccaactgacactgaggttcagatgcaaatggaagaaag 522  
DB 462 caggttagtgctagctcagctgattggtcaaatggaatgcaagcagatgtagattg 521  
QY 523 aaactggaagaaacacttgcatcttggtgtgtacatcatgcatcttgcaaaactta 582  
DB 522 gtgagccaagcgctgaacgcgaggtactcggaacatcgaggttttaccaaatccta 581  
QY 583 gctgaaggaagaaatcagaggtcttggtgcaaggtggttaccatatacaagaagca 642  
DB 582 caatcagaagagtaaaaggtltalgaaggtgttcttccaaacatccgagagcattt 641  
QY 643 ctgttgatattggaagatttaacacactatgatacagtggaacactcttggtattgaa 702  
DB 642 ctatgataatggaagactagcttgctatgatacagccaacacttgcatacgaataag 701  
QY 703 acacacttgagacaataatcatgactacaggtttatcaagttatgttctgagctgta 762  
DB 702 aagattgtgagataaactatttgcgcaactcttgctctataatgtagtggtctgtct 761  
QY 763 gctctattcttggaacacagccgagatgtcatcaaaagcgaataatgatacaaccca 822  
DB 762 tcgacaagatttgatgtgccaagctgagtggtgtgtaagcagagatgataacccaggtg 821  
QY 823 gataacaagaagggagctttgtataaatcatcagctacagctgattcagagctgtt 882

PR	03-JUN-1999	9905-0137528
PR	04-JUN-1999	9905-0137752
PR	07-JUN-1999	9905-0137720
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	10-JUN-1999	9905-0138847
PR	14-JUN-1999	9905-0139119
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139453
PR	17-JUN-1999	9905-0139460
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139459
PR	18-JUN-1999	9905-0139460
PR	18-JUN-1999	9905-0139461
PR	18-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0139463
PR	18-JUN-1999	9905-0139750
PR	18-JUN-1999	9905-0139763
PR	21-JUN-1999	9905-0139617
PR	22-JUN-1999	9905-0139699
PR	23-JUN-1999	9905-0140353
PR	23-JUN-1999	9905-0140354
PR	24-JUN-1999	9905-0140695
PR	28-JUN-1999	9905-0140823
PR	29-JUN-1999	9905-0140991
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141842
PR	01-JUL-1999	9905-0142154
PR	01-JUL-1999	9905-0142055
PR	06-JUL-1999	9905-0142390
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142820
PR	12-JUL-1999	9905-0142937
PR	13-JUL-1999	9905-0143342
PR	14-JUL-1999	9905-0143624
PR	15-JUL-1999	9905-0144005
PR	16-JUL-1999	9905-0144085
PR	16-JUL-1999	9905-0144086
PR	19-JUL-1999	9905-0144325
PR	20-JUL-1999	9905-0144332
PR	20-JUL-1999	9905-0144384
PR	21-JUL-1999	9905-0144614
PR	21-JUL-1999	9905-0145086
PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145085
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145089
PR	23-JUL-1999	9905-0145152
PR	23-JUL-1999	9905-0145184
PR	23-JUL-1999	9905-0145218
PR	23-JUL-1999	9905-0145224
PR	26-JUL-1999	9905-0145576
PR	27-JUL-1999	9905-0145513
PR	27-JUL-1999	9905-0145518
PR	28-JUL-1999	9905-0145519
PR	28-JUL-1999	9905-0145521
PR	02-AUG-1999	9905-0146586
PR	02-AUG-1999	9905-0146588
PR	02-AUG-1999	9905-0146589
PR	03-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147502



PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 11.9%; Score 124; DB 21; Length 1242;

Best Local Similarity 53.5%; Pred. No. 2,9e-30;  
 Matches 293; Conservative 0; Mismatches 240; Indels 15; Caps 1;  
 QY 463 cagtttttagccatccacactgactagtgaaagttcaagatgaacaaaggaagaaag 522  
 Db 464 caggttagtggctacgcagctgactgtaagtgaaaggaagaaaggaagatg 523  
 QY 523 aaactggaaggaacacacttgcatttcgtgtacatcatgcatgtaaatctta 582  
 Db 524 gtgagccaaagcccgaaacccgagttactcggaccatcgaggtctttaccaaatccta 583  
 QY 583 gctgaaggaagaaacacgagggcttggcaggtctgggtaccacaaatacaagagcga 642  
 Db 584 caatcagaaggaagaaaggttatgaaaggtgttctccaacatccagagagcat 643  
 QY 643 ctgtgtataatggaagatttaacacactatgatacagtgaaacactctgtatgaat 702  
 Db 644 ctagtgaataatggagaaactagctgtctatgacgcacaaacacttgcatacgaatgaag 703  
 QY 703 acaccacttgaggaacaaatcatgactacaggtttatcaagttatgtctgagactgta 762  
 Db 704 aagattgctgaggaataacatcttgcgcacactctgtcttataatgctccggtctgct 763  
 QY 763 gctctatctcgggaacacccagccgcatgtcatcaaaagcagaataatgaatcaacacga 822  
 Db 764 tcgacaagtttgagttgctcagcagtgatgtgtgaagaagatgagatgaacacaggtgaa 823  
 QY 823 gataaacaaggaagggacttctgtataatcatcagactgactctgtatcaagctgtt 882  
 Db 824 aat-----gctgtgtacagaatcttaccagactgtgttggatgaagcgtt 868  
 QY 883 caaggtgaaggaatcatgagctctatataaagcttttaccactctgtctgagatgacc 942  
 Db 869 aagtttgaaggaatacaagagcttctggaaggttctcccgacaatgggcaaggtcttga 928  
 QY 943 cctgtgcaatggtgtctgcttactactataaanaaatcaagagatgagtgagtcagt 1002  
 Db 929 ccgtggcaggtcgtgttctgtgtctccatgagaagtttaagcagctgcaggaatcct 988  
 QY 1003 ccaattta 1010  
 Db 989 tccttcta 996  
 RESULT 10  
 C46816  
 ID C46816 standard; DNA; 1103 BP.  
 XX  
 AC C46816:  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 51508.  
 XX  
 KW Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway;  
 KW metabolic pathway; promoter; termination sequence; ss.  
 OS Arabidopsis thaliana.  
 PN EPI033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 FE 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-01332485.  
PR 06-MAY-1999; 99US-01332486.  
PR 06-MAY-1999; 99US-0133487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.

PR 13-OCT-1999: 99US-0159294.  
PR 14-OCT-1999: 99US-0159295.  
PR 14-OCT-1999: 99US-0159329.  
PR 14-OCT-1999: 99US-0159330.  
PR 14-OCT-1999: 99US-0159331.  
PR 14-OCT-1999: 99US-0159637.  
PR 14-OCT-1999: 99US-0159638.  
PR 18-OCT-1999: 99US-0159584.  
PR 21-OCT-1999: 99US-0160741.  
PR 21-OCT-1999: 99US-0160767.  
PR 21-OCT-1999: 99US-0160768.  
PR 21-OCT-1999: 99US-0160770.  
PR 21-OCT-1999: 99US-0160814.  
PR 21-OCT-1999: 99US-0160815.  
PR 22-OCT-1999: 99US-0160880.  
PR 22-OCT-1999: 99US-0160981.  
PR 22-OCT-1999: 99US-0160989.  
PR 25-OCT-1999: 99US-0161404.  
PR 25-OCT-1999: 99US-0161405.  
PR 25-OCT-1999: 99US-0161406.  
PR 26-OCT-1999: 99US-0161359.  
PR 26-OCT-1999: 99US-0161360.  
PR 26-OCT-1999: 99US-0161361.  
PR 28-OCT-1999: 99US-0161920.  
PR 28-OCT-1999: 99US-0161992.  
PR 28-OCT-1999: 99US-0161993.  
PR 29-OCT-1999: 99US-0162142.

Query Match 10.9%; Score 113.4; DB 21; Length 1103;  
Best Local Similarity 50.6%; Pred. No. 8.3e-27;  
Matches 301; Conservative 0; Mismatches 291; Indels 3; Gaps 1;

QY 241 tataagggaatggttgcgcacagcccttagggatcatatgaagaaggcttctaagctt 300  
DB 345 taltcgaagattgttgggaacattgtgtacatagcaagggaagggttcttacta 404  
QY 301 tggcaagggaatgaccccgccattacagacagtagtattcttggagggtcgaaatg 360  
DB 405 tggaaagggttcttactctgattgcacgtcgaatgtcttcttggagggttcttgg 464  
QY 361 acatatgaacatctccgagaggttgttggcaaaagtgaagtgcattccctt 420  
DB 465 atgtatagcggttgaanaactgtatgttggaaagactttagtggatggtccatg 524  
QY 421 tggaaatcaatcttggaggaatgagtggtgttcttggcaagttttagccaatcca 480  
DB 525 agcaagaanaattcttctgtgttgcacacaggtgcactggtatcatgtlgaacaa 584  
QY 481 actgaacctagtgaaagttcagatgcagaaatgaaagaaagaaactggaagaaac 540  
DB 585 actgaaccttggaaagttcagatgcagaaagaaagaaagaaagaaagaaagaa 641  
QY 541 ttgcgatttcgtgtgtacatcalttgcataaaactttagctgaaggagaatacga 600  
DB 642 agacggtattctggagcgctgaatgtcttacaacttggagacaggaagatcgca 701  
QY 601 gggccttggcagcgctgggaaccaatacaagacagcactgtggaataggagat 660  
DB 702 gctcttggactgtcttggcactaagctagcaagaacgtatatacattgtcgtcga 761  
QY 661 ttaccacttatagacaggaacacttacttgatgtatgaatacaccacttggagaa 720  
DB 762 tttagcgattacagatacaggaagagacatcttgaagttccgggttcaactgaac 821  
QY 721 ataatgaatcaggttatcaagttatgttctgactgttgaatcttattcttgagaca 780  
DB 822 gtctgcacacataatctatctgactgtggagcaggaattcttctgtcgtcatggtct 881  
QY 781 ccagcgatgtcatcaaaagcaataatgaatcaaccacagagataaacaaggaa 835  
DB 882 cctgtcagcgtgttgaatgaatgaatgaatgtggagattctgtgttacaaggaca 936

RESULT 11  
C33096  
ID C33096 standard; DNA; 1227 BP.  
XX  
AC C33096;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 1765.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KM protein identification; signal transduction pathway;  
KM metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.

PR	18-JUN-1999;	99US-0139459;	PR	20-AUG-1999;	99US-0149929;
PR	18-JUN-1999;	99US-0139460;	PR	23-AUG-1999;	99US-0149902;
PR	18-JUN-1999;	99US-0139461;	PR	23-AUG-1999;	99US-0149930;
PR	18-JUN-1999;	99US-0139462;	PR	25-AUG-1999;	99US-0150566;
PR	18-JUN-1999;	99US-0139463;	PR	26-AUG-1999;	99US-0150884;
PR	18-JUN-1999;	99US-0139750;	PR	27-AUG-1999;	99US-0151065;
PR	21-JUN-1999;	99US-0139817;	PR	27-AUG-1999;	99US-0151066;
PR	22-JUN-1999;	99US-0139899;	PR	27-AUG-1999;	99US-0151080;
PR	23-JUN-1999;	99US-0140353;	PR	30-AUG-1999;	99US-0151303;
PR	24-JUN-1999;	99US-0140695;	PR	31-AUG-1999;	99US-0151438;
PR	28-JUN-1999;	99US-0140823;	PR	01-SEP-1999;	99US-0151930;
PR	29-JUN-1999;	99US-0140991;	PR	07-SEP-1999;	99US-0152363;
PR	30-JUN-1999;	99US-0141287;	PR	10-SEP-1999;	99US-0153070;
PR	01-JUL-1999;	99US-0141842;	PR	13-SEP-1999;	99US-0153758;
PR	01-JUL-1999;	99US-0142154;	PR	15-SEP-1999;	99US-0154018;
PR	02-JUL-1999;	99US-0142055;	PR	16-SEP-1999;	99US-0154039;
PR	06-JUL-1999;	99US-0142380;	PR	20-SEP-1999;	99US-0154779;
PR	08-JUL-1999;	99US-0142803;	PR	22-SEP-1999;	99US-0155139;
PR	09-JUL-1999;	99US-0142920;	PR	23-SEP-1999;	99US-0155486;
PR	12-JUL-1999;	99US-0142977;	PR	24-SEP-1999;	99US-0155659;
PR	13-JUL-1999;	99US-0143542;	PR	28-SEP-1999;	99US-0156458;
PR	14-JUL-1999;	99US-0143624;	PR	29-SEP-1999;	99US-0156556;
PR	15-JUL-1999;	99US-0144005;	PR	04-OCT-1999;	99US-0157117;
PR	16-JUL-1999;	99US-0144085;	PR	05-OCT-1999;	99US-0157753;
PR	16-JUL-1999;	99US-0144086;	PR	06-OCT-1999;	99US-0157865;
PR	19-JUL-1999;	99US-0144325;	PR	07-OCT-1999;	99US-0158029;
PR	19-JUL-1999;	99US-0144331;	PR	08-OCT-1999;	99US-0158232;
PR	19-JUL-1999;	99US-0144332;	PR	12-OCT-1999;	99US-0158369;
PR	19-JUL-1999;	99US-0144333;	PR	13-OCT-1999;	99US-0159293;
PR	19-JUL-1999;	99US-0144334;	PR	13-OCT-1999;	99US-0159295;
PR	19-JUL-1999;	99US-0144335;	PR	14-OCT-1999;	99US-0159329;
PR	20-JUL-1999;	99US-0144352;	PR	14-OCT-1999;	99US-0159330;
PR	20-JUL-1999;	99US-0144632;	PR	14-OCT-1999;	99US-0159331;
PR	21-JUL-1999;	99US-0144884;	PR	14-OCT-1999;	99US-0159637;
PR	21-JUL-1999;	99US-0145086;	PR	18-OCT-1999;	99US-0159638;
PR	21-JUL-1999;	99US-0145088;	PR	21-OCT-1999;	99US-0160741;
PR	22-JUL-1999;	99US-0145085;	PR	21-OCT-1999;	99US-0160767;
PR	22-JUL-1999;	99US-0145089;	PR	21-OCT-1999;	99US-0160768;
PR	22-JUL-1999;	99US-0145192;	PR	21-OCT-1999;	99US-0160770;
PR	23-JUL-1999;	99US-0145145;	PR	21-OCT-1999;	99US-0160814;
PR	23-JUL-1999;	99US-0145218;	PR	22-OCT-1999;	99US-0160980;
PR	23-JUL-1999;	99US-0145224;	PR	22-OCT-1999;	99US-0160981;
PR	26-JUL-1999;	99US-0145276;	PR	22-OCT-1999;	99US-0160989;
PR	27-JUL-1999;	99US-0145913;	PR	25-OCT-1999;	99US-0161404;
PR	27-JUL-1999;	99US-0145919;	PR	25-OCT-1999;	99US-0161405;
PR	28-JUL-1999;	99US-0145951;	PR	26-OCT-1999;	99US-0161359;
PR	02-AUG-1999;	99US-0146386;	PR	26-OCT-1999;	99US-0161360;
PR	02-AUG-1999;	99US-0146388;	PR	26-OCT-1999;	99US-0161361;
PR	03-AUG-1999;	99US-0147038;	PR	28-OCT-1999;	99US-0161920;
PR	04-AUG-1999;	99US-0147204;	PR	28-OCT-1999;	99US-0161992;
PR	04-AUG-1999;	99US-0147302;	PR	28-OCT-1999;	99US-0161993;
PR	05-AUG-1999;	99US-0147192;	PR	29-OCT-1999;	99US-0162142;
PR	05-AUG-1999;	99US-0147260;			
PR	06-AUG-1999;	99US-0147303;			
PR	06-AUG-1999;	99US-0147416;			

PR	11-MAY-1999;	9905-0134256;
PR	14-MAY-1999;	9905-0134218;
PR	14-MAY-1999;	9905-0134219;
PR	14-MAY-1999;	9905-0134221;
PR	14-MAY-1999;	9905-0134270;
PR	18-MAY-1999;	9905-0134768;
PR	19-MAY-1999;	9905-0134641;
PR	20-MAY-1999;	9905-0135124;
PR	21-MAY-1999;	9905-0135553;
PR	24-MAY-1999;	9905-0135629;
PR	27-MAY-1999;	9905-0136021;
PR	28-MAY-1999;	9905-0136592;
PR	01-JUN-1999;	9905-0136782;
PR	03-JUN-1999;	9905-0137222;
PR	04-JUN-1999;	9905-0137528;
PR	04-JUN-1999;	9905-0137502;
PR	08-JUN-1999;	9905-0137724;
PR	10-JUN-1999;	9905-0138094;
PR	10-JUN-1999;	9905-0138540;
PR	14-JUN-1999;	9905-0138847;
PR	14-JUN-1999;	9905-0139119;
PR	16-JUN-1999;	9905-0139452;
PR	17-JUN-1999;	9905-0139493;
PR	18-JUN-1999;	9905-0139452;
PR	18-JUN-1999;	9905-0139454;
PR	18-JUN-1999;	9905-0139455;
PR	18-JUN-1999;	9905-0139457;
PR	18-JUN-1999;	9905-0139457;
PR	18-JUN-1999;	9905-0139458;
PR	18-JUN-1999;	9905-0139459;
PR	18-JUN-1999;	9905-0139750;
PR	18-JUN-1999;	9905-0139763;
PR	21-JUN-1999;	9905-0139617;
PR	22-JUN-1999;	9905-0139699;
PR	23-JUN-1999;	9905-0140353;
PR	23-JUN-1999;	9905-0140354;
PR	24-JUN-1999;	9905-0140695;
PR	26-JUN-1999;	9905-0140823;
PR	29-JUN-1999;	9905-0140991;
PR	30-JUN-1999;	9905-0141287;
PR	01-JUL-1999;	9905-0141842;
PR	01-JUL-1999;	9905-0142154;
PR	02-JUL-1999;	9905-0142055;
PR	06-JUL-1999;	9905-0142390;
PR	08-JUL-1999;	9905-0142803;
PR	09-JUL-1999;	9905-0142927;
PR	12-JUL-1999;	9905-0143277;
PR	13-JUL-1999;	9905-0143542;
PR	14-JUL-1999;	9905-0143624;
PR	15-JUL-1999;	9905-0144005;
PR	16-JUL-1999;	9905-0144086;
PR	19-JUL-1999;	9905-0144325;
PR	19-JUL-1999;	9905-0144325;
PR	19-JUL-1999;	9905-0144331;
PR	19-JUL-1999;	9905-0144332;
PR	19-JUL-1999;	9905-0144333;
PR	19-JUL-1999;	9905-0144334;
PR	19-JUL-1999;	9905-0144335;
PR	20-JUL-1999;	9905-0144352;
PR	20-JUL-1999;	9905-0144632;
PR	20-JUL-1999;	9905-0144684;
PR	21-JUL-1999;	9905-0144814;
PR	21-JUL-1999;	9905-0145086;
PR	21-JUL-1999;	9905-0145088;
PR	22-JUL-1999;	9905-0145085;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145192;
PR	23-JUL-1999;	9905-0145193;



XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 9905-0121825.  
PR 05-MAR-1999; 9905-0123180.  
PR 09-MAR-1999; 9905-0123548.  
PR 23-MAR-1999; 9905-0125788.  
PR 25-MAR-1999; 9905-0126264.  
PR 29-MAR-1999; 9905-0126785.  
PR 01-APR-1999; 9905-0127462.  
PR 06-APR-1999; 9905-0128234.  
PR 08-APR-1999; 9905-0128714.  
PR 16-APR-1999; 9905-0129645.  
PR 19-APR-1999; 9905-0130077.  
PR 21-APR-1999; 9905-0130049.  
PR 23-APR-1999; 9905-0130510.  
PR 28-APR-1999; 9905-0130891.  
PR 30-APR-1999; 9905-0131449.  
PR 30-APR-1999; 9905-0132048.  
PR 04-MAY-1999; 9905-0132407.  
PR 05-MAY-1999; 9905-0132485.  
PR 06-MAY-1999; 9905-0132486.  
PR 07-MAY-1999; 9905-0132487.  
PR 07-MAY-1999; 9905-0132653.  
PR 11-MAY-1999; 9905-0134256.  
PR 14-MAY-1999; 9905-0134218.  
PR 14-MAY-1999; 9905-0134219.  
PR 14-MAY-1999; 9905-0134221.  
PR 14-MAY-1999; 9905-0134370.  
PR 18-MAY-1999; 9905-0134768.  
PR 19-MAY-1999; 9905-0134941.  
PR 20-MAY-1999; 9905-0135124.  
PR 21-MAY-1999; 9905-0135353.  
PR 24-MAY-1999; 9905-0135629.  
PR 25-MAY-1999; 9905-0136021.  
PR 27-MAY-1999; 9905-0136392.  
PR 28-MAY-1999; 9905-0136782.  
PR 03-JUN-1999; 9905-0137222.  
PR 03-JUN-1999; 9905-0137528.  
PR 04-JUN-1999; 9905-0137502.  
PR 07-JUN-1999; 9905-0137724.  
PR 08-JUN-1999; 9905-0138094.  
PR 10-JUN-1999; 9905-0138540.  
PR 10-JUN-1999; 9905-0138847.  
PR 14-JUN-1999; 9905-0139119.  
PR 16-JUN-1999; 9905-0139452.  
PR 16-JUN-1999; 9905-0139453.  
PR 17-JUN-1999; 9905-0139492.  
PR 18-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139459.  
PR 18-JUN-1999; 9905-0139460.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 18-JUN-1999; 9905-0139750.  
PR 18-JUN-1999; 9905-0139763.  
PR 21-JUN-1999; 9905-0139817.  
PR 22-JUN-1999; 9905-0139899.  
PR 23-JUN-1999; 9905-0140353.  
PR 23-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140695.  
PR 28-JUN-1999; 9905-0140823.  
PR 29-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.

PR 01-JUL-1999; 9905-0141842.  
PR 01-JUL-1999; 9905-0142154.  
PR 02-JUL-1999; 9905-0142055.  
PR 06-JUL-1999; 9905-0142390.  
PR 08-JUL-1999; 9905-0142803.  
PR 09-JUL-1999; 9905-0142920.  
PR 12-JUL-1999; 9905-0142977.  
PR 13-JUL-1999; 9905-0143542.  
PR 14-JUL-1999; 9905-0143624.  
PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144333.  
PR 19-JUL-1999; 9905-0144334.  
PR 19-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144632.  
PR 20-JUL-1999; 9905-0144684.  
PR 21-JUL-1999; 9905-0144814.  
PR 21-JUL-1999; 9905-0145086.  
PR 21-JUL-1999; 9905-0145088.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 22-JUL-1999; 9905-0145192.  
PR 23-JUL-1999; 9905-0145145.  
PR 23-JUL-1999; 9905-0145218.  
PR 23-JUL-1999; 9905-0145224.  
PR 26-JUL-1999; 9905-0145276.  
PR 27-JUL-1999; 9905-0145913.  
PR 27-JUL-1999; 9905-0145918.  
PR 28-JUL-1999; 9905-0145951.  
PR 02-AUG-1999; 9905-0146386.  
PR 02-AUG-1999; 9905-0146388.  
PR 02-AUG-1999; 9905-0146389.  
PR 03-AUG-1999; 9905-0147038.  
PR 04-AUG-1999; 9905-0147204.  
PR 04-AUG-1999; 9905-0147302.  
PR 05-AUG-1999; 9905-0147192.  
PR 05-AUG-1999; 9905-0147260.  
PR 06-AUG-1999; 9905-0147303.  
PR 06-AUG-1999; 9905-0147416.  
PR 09-AUG-1999; 9905-0147493.  
PR 09-AUG-1999; 9905-0147935.  
PR 10-AUG-1999; 9905-0148171.  
PR 11-AUG-1999; 9905-0148319.  
PR 12-AUG-1999; 9905-0148341.  
PR 13-AUG-1999; 9905-0148565.  
PR 13-AUG-1999; 9905-0148684.  
PR 16-AUG-1999; 9905-0149368.  
PR 17-AUG-1999; 9905-0149175.  
PR 18-AUG-1999; 9905-0149426.  
PR 20-AUG-1999; 9905-0149722.  
PR 20-AUG-1999; 9905-0149723.  
PR 20-AUG-1999; 9905-0149929.  
PR 23-AUG-1999; 9905-0149902.  
PR 23-AUG-1999; 9905-0149930.  
PR 25-AUG-1999; 9905-0150566.  
PR 26-AUG-1999; 9905-0150884.  
PR 26-AUG-1999; 9905-0151065.  
PR 27-AUG-1999; 9905-0151066.  
PR 27-AUG-1999; 9905-0151080.  
PR 30-AUG-1999; 9905-0151303.  
PR 31-AUG-1999; 9905-0151438.  
PR 01-SEP-1999; 9905-0151930.  
PR 07-SEP-1999; 9905-0152363.  
PR 10-SEP-1999; 9905-0153070.  
PR 13-SEP-1999; 9905-0153758.  
PR 15-SEP-1999; 9905-0154018.





CC against NHP and in assays to identify modulators (agonists and  
CC antagonists) of NHP expression and activity. Anti-NHP antibodies and  
CC NHP antagonists may also be used to down regulate NHP expression and  
CC activity. Anti-NHP antibodies may also be used as diagnostic agents for  
CC detecting the presence of NHP polypeptides in samples. Processes and  
CC conditions associated with NHP expression and activity include  
CC thermogenesis, obesity and cachexia.

XX  
XX  
XX  
Sequence 876 BP; 237 A; 184 C; 229 G; 226 T; 0 other:

Query Match 9.0%; Score 93.4; DB 21; Length 876;  
Best Local Similarity 48.3%; Pred. No. 2,7e-20;  
Matches 419; Conservative 0; Mismatches 406; Indels 42; Gaps 4;

QY 124 ggcgtacacgtgacgagctgacgaacattcccttgatctcaaaaaactgcactca 183  
DB 43 gctccatcactcgtgagtcggtacattcccaattgtaacaaagacagctccag 102  
QY 184 atgcaagagagaagcagctcttgcgttgagagagcgtgcaagaagatctgccccat 243  
DB 103 attcaagagcgaagaatgatgcaaaattgaagaattaga-----tac 147  
QY 244 aggggaatgtgacgacagccctagagatcatgaaagagagccttctaagcttgcg 303  
DB 148 cgaaggaatgtgacgacgattagtagagatagcagaagaagagcgtgaagacacttac 207  
QY 304 caaggaatgacacccgcaattacagacagtagtattcttgaagtcgaatgacaca 363  
DB 208 tcggagatgccccgcgagatgttacgcagcatccatgacacatagaatagcact 267  
QY 364 taigaacatcccgagaggtgtgtgttgcaaaagtgaagtgcattaccctcttgcg 423  
DB 268 taacagagcttgaagcgaactatcatgaaagcccaagaagatg-----aaccttacg 321  
QY 424 aaatcagtcattggaaggtatgctggtgttatttgccagctttagccaatcaact 483  
DB 322 ataattgtagatgtggaattctgtctgagctatctatcttaacacatgccaatcaact 381  
QY 484 gaactatggaagtctcagaatgaatgaagaagaagaactgcgaagaactatgc 543  
DB 382 gatgtttgaatgaatcgaatgacgaagaacacacattcaagaagagaa----- 433  
QY 544 cgatttggtgtgacatgacatttgcgaagaacttgaatgagagaggaatcgaagg 603  
DB 434 -----tgataggaacttcag-----aacattacagcagaagaggaagaaga 480  
QY 604 ctctggcagagctgtggtlaaccaataaagaagcagcagctgtgaatattggaattta 663  
DB 481 ctgtggaagaggtgtgtcccttactgcgagagaggtgtgtgtgtgtgtgtgtgtgt 540  
QY 664 accaatatgacagtgaaacactactgtgtattgaatcacacacttgagacaatc 723  
DB 541 ccggtlcatgacatcacaaagaacatttactctcgaagccgcatgtgagacactgtg 600  
QY 724 atgctacaggttatcaagttatgttctgactgtgagcttattcttctggaacacca 783  
DB 601 tataccaacttctcgaagcttcaactgtgtctggaagggccctggtcctcaaacct 660  
QY 784 ggcgaatgcatcaaaaagacaataatgaatcaacacagagaataaacaagaagggact 843  
DB 661 gtgtatgtgtgtgagacagcgtatgataacagagaggtctgagatgagatgttct 720  
QY 844 ttgataaatcatcgaactgactgtgtattcaagctgttcaaggtgaagattcaatgct 903  
DB 721 ggcatacagaagaccctgtgattgtgttacaagatgaaagtgaaggttttttgc 780  
QY 904 ctataaagagcttttaccacttgcctgtgagatgagccctgtgcaatgtgttcgg 963  
DB 781 ctctataaagcttttggccaattgtgtgagacttgccttgaatataattttctt 840  
QY 964 ctactatgaaaaaatcagaagagatg 990

DB 841 gtgacatagcagcttgaagaactg 867

RESULT 15

A27999

ID A27999 standard; DNA; 1031 BP.

AC A27999;

DT 29-AUG-2000 (first entry)

DE Murine uncoupling protein isoform mUCP5 nucleotide sequence.

KW Uncoupling protein 5; UCP5; metabolism; H+ leak; mouse; ds; infection;  
metabolic rate; obesity; stroke; trauma; burn trauma; sepsis.

OS Mus musculus.

FT Key Location/Qualifiers

FT CDS 28..996

FT /\*tag= a

FT /product= "UCP5"

FT /note= "Uncoupling protein 5"

FT /transl\_except= (pos:403..405, aa:Gly)

PN MO200032624-A2.

PD 08-JUN-2000.

PF 03-NOV-1999; 99MO-US25947.

PR 30-NOV-1998; 98US-0110286.

PR 16-APR-1999; 99US-0129583.

PR 15-JUL-1999; 99US-0143886.

PA (GERTH) GENENTECH INC.

PI Adams S, Pan J;

DR WPI: 2000-412284/35.

DR P-PSDB: Y94669.

PT Isolated nucleic acid encodes human uncoupling protein 5 useful in  
diagnostic assays and treatment of obesity, stroke, trauma, sepsis and  
infection.

PS Claim 3; Fig 15; 90pp; English.

This sequence represents a murine uncoupling protein 5 isoform mUCP5  
encoding DNA sequence. UCP5 is involved in metabolism, and it may be  
involved in catalyzing H+ leak, and therefore be involved in energetic  
inefficiency in vivo. The present invention relates to human and murine  
UCP5 nucleotide and protein sequences. There are three isoforms of human  
UCP5, hUCP5L, hUCP5S1, and two isoforms of murine UCP5, mUCP5L  
and mUCP5S. The human UCP5 gene is located on chromosome 10q23-25. The  
nucleic acids encoding UCP5 can be used as hybridization probes, in  
chromosome and gene mapping, for the generation of antisense RNA and DNA  
and in the preparation of recombinant UCP5 proteins. UCP5 nucleic acids  
can be used in gene therapy for regulation of metabolic conditions.  
Upregulating or downregulating UCP5 activity in a mammal is used for  
modulating metabolic rate in the mammal, in particular upregulation of  
UCP5 activity stimulates an increase in metabolic rate in an obese  
mammal. Other therapeutic applications associated with modulating UCP5  
activity are treating symptoms associated with stroke, trauma (e.g. burn  
trauma), sepsis and infection. Detecting UCP5 activity can be used to  
assist predictions concerning metabolic conditions or risk for onset of  
obesity and as UCP5 may control the generation of reactive oxygen to  
diagnose impaired neural activity or neural degeneration. Anti-UCP5  
antibodies can be used in diagnostic assays and for the affinity  
purification of UCP5 from recombinant cell culture or natural sources.  
Sequence 1031 BP; 270 A; 199 C; 264 G; 298 T; 0 other:

Query Match 8.6%; Score 89.6; DB 21; Length 1031;  
 Best Local Similarity 48.2%; Pred. No. 5, 2e-19;  
 Matches 417; Conservative 0; Mismatches 404; Indels 45; Gaps 4;

```

QY 117 cggctgcgagcgtacccggtgacagacaccccttgatcaccacaaactcg 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 cggccttgctcctatgtgttcgagctgcgagcttccttgatcttacttaaacacg 227
QY 177 actccaatgcagagagagagagagagagagagagagagagagagagagagag 236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 gcttgcaagtcacagagagagagagagagagagagagagagagagagagagag 276
QY 237 cccctataagggagagagagagagagagagagagagagagagagagagagag 296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 ----tatagagagatgtttcattgctgttcggaatcctaaagagagagagatctg 332
QY 297 gcttggcagagagagagagagagagagagagagagagagagagagagagag 356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 333 tctgtatcagagagagagagagagagagagagagagagagagagagagagag 392
QY 357 ggtcacatagagagagagagagagagagagagagagagagagagagagagag 416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 393 tggatattatgaaagcttgaagcattatgttagaagcttggaaagagagagagctt 452
QY 417 ccttggaaatcagatcattggagagagagagagagagagagagagagagagag 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 453 aattaca-----tgatctgaggagagagagagagagagagagagagagagag 506
QY 477 tccaactgagcagagagagagagagagagagagagagagagagagagagagag 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 507 tcccaactgagcagagagagagagagagagagagagagagagagagagagagag 566
QY 537 accatlgcagatctggtggtacatcattgcaaaaatcctagctgaagagagagat 596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 567 gattggcagcttcattgacatatacca-----gcaagaaggtac 605
QY 597 acgagggcttggcagagctgggtaccacatatacaaaagagagagagagagagag 656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 606 caggggctcgtgagagagagagagagagagagagagagagagagagagagagag 665
QY 657 agatttaaccacttaagatacagtgaaacactactgtgattgaatacaccacttgaga 716
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 666 agagctgcccgttattgattacacagagagagagagagagagagagagagagag 725
QY 717 caatatcagactcagaggttatacaagttatgcttgacagtgagcttctatctggg 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 726 caaatlttaacacacttgttccagtttcaactggtggttggctggggtcttgacac 785
QY 777 aacaccagcagatgtatcaaaaagagagagagagagagagagagagagagagagag 836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 786 taacctgtgagatgtgagagagagagagagagagagagagagagagagagagagag 842
QY 837 gggacttctgtataatcagactgagctgtgattcagagctgttcaaggtgaagagat 896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 843 tgtgaccttacaagagagagagagagagagagagagagagagagagagagagag 902
QY 897 catgagctcattaaagagagagagagagagagagagagagagagagagagagagag 956
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 903 ttctgacctctaaagagagagagagagagagagagagagagagagagagagagag 962
QY 957 gttctgacttaactatgaaataca 982
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 963 ttttttattacattagcagctca 988
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2001, 17:33:01 ; Search time 66.09 seconds  
(without alignments)  
2744.927 Million cell updates/sec

Title: US-09-397-342-2

Perfect score: 1039

Sequence: 1 ccgagctcgcgtcccgcttat.....cagatatccatcacctgyc 1039

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 302621 seqs, 87301344 residues

tal number of hits satisfying chosen parameters: 605242

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/prodata1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata1/ina/5A\_COMB.seq:\*  
4: /cgn2\_6/prodata1/ina/5B\_COMB.seq:\*  
5: /cgn2\_6/prodata1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata1/ina/backfillsl.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50.8	4.9	1949	2	US-08-937-466-3
2	50.8	4.9	1949	2	US-09-172-528-3
3	50.8	4.9	1949	3	US-09-318-199-3
4	49.4	4.8	1255	1	US-08-518-878B-38
5	49.4	4.8	1255	1	US-08-294-522B-38
6	49.4	4.8	1255	2	US-08-470-868A-38
7	49.4	4.8	1536	2	US-08-807-861A-38
8	49.4	4.8	1536	3	US-09-210-681-38
9	49.4	4.8	1536	4	US-08-946-719A-38
10	49.2	4.7	2782	2	US-08-937-466-1
11	49.2	4.7	2782	2	US-09-172-528-1
12	49.2	4.7	2782	3	US-09-318-199-1
13	48.6	4.7	1777	2	US-08-937-466-5
14	48.6	4.7	1777	2	US-09-172-528-5
15	48.6	4.7	1777	3	US-09-318-199-5
16	45.4	4.4	7218	1	US-08-232-463-14
17	37	3.6	1192	4	US-09-142-565-1
18	34	3.3	1205	1	US-08-518-878B-36
19	34	3.3	1205	1	US-08-294-522B-37
20	34	3.3	1205	2	US-08-807-861A-36
21	34	3.3	1205	3	US-08-470-868A-36
22	34	3.3	1205	3	US-09-210-681-36
23	34	3.3	1205	4	US-08-946-719A-36
24	31	3.0	3931	1	US-07-999-280A-21
25	31	3.0	3931	1	US-08-426-036-21
26	31	3.0	3931	1	US-08-426-279-21
27	31	3.0	3931	1	US-08-401-013-21

28	31	3.0	3931	4	US-08-426-570-21	Sequence 21, Appl
29	31	3.0	3931	4	US-08-425-876-21	Sequence 21, Appl
30	31	3.0	3931	4	US-08-426-243-21	Sequence 21, Appl
31	31	3.0	3931	4	US-08-401-632-21	Sequence 21, Appl
32	31	3.0	4977	1	US-08-030-096-7	Sequence 7, Appl
33	30.8	3.0	7617	3	US-08-646-538-34	Sequence 34, Appl
34	30.6	2.9	847	4	US-09-142-565-5	Sequence 5, Appl
35	30	2.9	1882	1	US-08-458-120-1	Sequence 1, Appl
36	30	2.9	1882	2	US-08-867-870-1	Sequence 1, Appl
37	30	2.9	1882	4	US-09-326-217-1	Sequence 1, Appl
38	30	2.9	2907	4	US-08-816-346-1	Sequence 1, Appl
39	30	2.9	2907	4	US-08-816-346-55	Sequence 55, Appl
40	30	2.9	2907	4	US-09-335-411-1	Sequence 1, Appl
41	30	2.9	2907	4	US-09-335-411-55	Sequence 55, Appl
42	30	2.9	51952	4	US-08-947-823-1	Sequence 1, Appl
43	29.8	2.9	1177	3	US-08-961-871-9	Sequence 9, Appl
44	29.8	2.9	1259	3	US-08-961-871-11	Sequence 11, Appl
45	29.8	2.9	3855	4	US-08-974-549A-4	Sequence 4, Appl

#### ALIGNMENTS

```
RESULT 1
US-08-937-466-3
; Sequence 3, Application US/08937466
; Patent No. 5846779
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: ucp3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,466
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1949 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-937-466-3

Query Match 4.9%; Score 50.8; DB 2; Length 1949;
Best Local Similarity 47.6%; Pred. No. 2.5e-07;
Matches 200; Conservative 0; Mismatches 202; Indels 18; Gaps 1;

OY 563 atgcatctgcaaaaacttgctgctgaaggaaggaactgagcagctgggtac 622
Db 718 ATGCTTACAGAACCATCGCCAGGAGGAGGAGGAGTGCAGCGCCTGTGGAAGGACTTGGC 777
```



```

;      TOPOLOGY: linear
;      MOLECULE TYPE: CDNA
US-09-318-199-3

```

Query Match	4.98;	Score 50.8;	DB 3;	Length 1949;
Best Local Similarity	47.68;	Pred. No. 2.5e-07;		

Matches 200; Conservative 0; Mismatches 202; Indels 18; Gaps 1,

QY	563	atgacattgcgaataactcttagccgaaaggaatacgaaggtcttggagcaactggtgac	622
Db	718	ATGCGCTACAGAACCATTCGCCAGGAGAGMAAGNACTAGGGGCGCTGTGTGAAGAGGACTTGGC	777
QY	623	ccaatatacaagaagcagcacctgylgtaatatggagattlaacccattatgatacagltga	682
Db	778	CCAACTACACAGAAATAGCCATTGTCAACTGTCTGAGATGGTGAACCTACGACATCA	837
QY	683	aacacactcttggatttgaataacacacacacttggagacaataatcatgactacaggttatcaa	742
Db	838	AGGGAAGATTGTCTGGAGCTTCACACTGTTTACTACAACTTCCCTGTCACTTTGTCTCTGTG	897
QY	743	gtttatgtcttcgaactgtagcttcatcttcttcgttgaaacacagccagatgcatcaaaaaga	802
Db	898	CCTTTGGAGCTGGGCTTTGTGTGCACACAGTGTGGCGCTCCCGGTGGATGTGGCTTAAGAACCC	957
QY	803	gaataatgaatacaaccacgagatlaaacaagaaggagacttltgtataataatcatgactg	862
Db	958	GATACATGAACGGCTCCCTTAGGCAAGGTACCGAGCCCTGTGC-----	999
QY	863	acgtgcttgatcagagctgltcaaggtgaaagatlaatgagtcataataaaagctltttaa	922
Db	1000	ACTGTATGCTGAAGATGTTGGCTCCAGAGAGGAGCCACGGCGCTTTCACAAACAGATTGTGGC	1058
QY	923	catcttgctcgagagatbaaccctctgttcaatagtgltctgggtctctataatgaaaaaatca	982
Db	1060	CTCTCCTTTCTCGCGCTGGAGACTTGGACGTGATGATGTTTGTAAACATATGTGCAACTGA	1119

RESULT 4  
US-08-518-878B-38

```

: GENERAL INFORMATION:
:
: APPLICANT: Tartaglia, Louis A.
:
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
:
: TITLE OF INVENTION: TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
:
: NUMBER OF SEQUENCES: 57
:
: CORRESPONDENCE ADDRESS:

```

STREET: 1155 Avenue of the Americas

CITY: New York  
STATE: New York  
COUNTRY: U.S.A.

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: DataOnTrk Release #1.0

SOFTWARE: PALCENLH release #1.0, version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/518,878E  
FILING DATE: 23-AUG-1995

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30 743

REGISTRATION NUMBER: 30,142  
REFERENCE/DOCKET NUMBER: 7853-036

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864  
TELEEX: 66141 PENNTE

INFORMATION FOR SEQ ID NO: 38:

```

; LENGTH: 1255 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-518-878B-38

```

Query Match	4.8%	Score	49.4	DB	1	Length	1255
Best Local Similarity	49.1%	Pred	No. 5.7e-07				
Matches	199	Conservative	0	Mismatches	191	Indels	15
						Gaps	2

QY 103 aaattccactacgtccggtgtgagcggtacacgltggccgagctcagaacattcccttgat 162  
Db 16 AAGTTCTTTGGGGCTGGGCACAGCTGCTCGATGCCAAATCATCATCCTTTCTCTGGAT 75  
QY 163 ctccacaaaacttcgactccaatgcgaaggaagaacgactctctgtcgtgttggaggaacggt 222  
Db 76 ACTGCTAAAGTCCGGTTACATCTCCAAAGGNAAG-----TCAGGGGCCAGTCCG 126  
QY 223 gcaagaagaactgcggccctatagggaatggtgcgacagccctaggatcatltaagag 282  
Db 127 GCTACAGTCACAGCCGCCAATGACCGGGGTGATGGGGCACCATTCTGACCAATGATGGCTACT 186  
QY 283 gaagcttcttcaagcttggcaaggagtgacaccgcgcatttcagaacgctagtgat 342  
Db 187 GAGGGCCCCCAAGCCTTCAATAGGGCTGTGTCCCGCCTCGACGCCCAATATAGCTTT 246  
QY 343 tctgtaggtcgcaatggtcacatatgaacatctccgagagtggtgtgttggcacaagtga 402  
Db 247 GCGCTGTGTCGGATGGCGCTGTATGATTTCTTCAACA-----GTTCTACACCAAGGC 300  
QY 403 gatgagcattatcccttggaaatcacgtaacttggaggatgaltgctgpbattatgac 462  
Db 301 TCTGAGCATGCGCAGATTGGGAGCGCGCTCTCTACAGCAGACACACAGGTGCTTGCT 360  
QY 463 cagttttaagccaatccaaactgacactagttgaagtttaagatga 507  
Db 361 GTGGCTGTGGCCCAAGCCACACGATGTGTAAAGTTCGATTCCA 405

## RESULT 5

; Sequence 38, Application US/08294522B  
: Patent No. 5741666

GENERAL INFORMATION:

APPLICANT: Tartaglia, Louis A.  
TITLE OF INVENTION: Composition

;	TITLE OF INVENTION:	Treatment of Body Weight Disorders, Including Obesity
:	NUMBER OF SEQUENCES:	48

CORRESPONDENCE ADDRESS:  
ADDRESS: Benito F.

STREET: 1155 Avenue of the Americas

;  
;  
; CITY: New York  
; STATE: New York

```

;      ZIP: 10036-2711
;      COMPUTER READABLE FORM:
;

```

MEDIUM TYPE: Floppy disk

COMPIER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
;

```
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/01  
FILING DATE: 23-AUG-1994

CLASSIFICATION: 514

ATTORNEI/AGENTI INFORMATION:  
NAME: Coruzzi, Laura A.

```

; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-015

```

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

; INFORMATION FOR SEQ ID NO:  
; SEQUENCE CHARACTERISTICS:









```

Db      958  LAATACATGAGCGCTCCCTTAGGCGAGGTACCGACGCCCTCTGC----- 999
OY      863  actgcttgatcagaagcgtgtcacaagtgaaagatcattgactatataaagccttllac 922
Db      1000  ACTGATATGCTGAAGATGAGCGGCTCAGAGAGGACCCACGCGCTTCACAAAGATTGTGTC 1059
OY      923  catcttgctgtgaatgatgacccttggtaatggtgtcttcggtctacttataaataatca 982
Db      1060  CCTCCTTTCTGCGCTGGAGAGCTTGCAACGATGATGTTGTAACATATGAGCAACTGA 1119

RESULT 12
US-09-318-199-1
; Sequence 1, Application US/09318199
; Patent No. 6025469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: ucp3 genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/318,199
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2782 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
;
US-09-318-199-1
Query Match 4.7%; Score 49.2; DB 3; Length 2782;
Best Local Similarity 47.4%; Pred. No. 1.2e-06;
Matches 199; Conservative 0; Mismatches 203; Indels 18; Gaps 1
OY      563  atgcatcttgcaaaaactctagctgaagagggatacgaagcgcttggcagcgctgggtac 622
Db      718  ATGCTTACGACGAAACATCCGCGCAGGAGGAGAGAGTCAAGGCGCCTGTGGAAGAGCACTTGGC 777
OY      623  ccaatacaaaagacagcagcactggtgaataatggagatttaaccaattatgatacagtga 682
Db      778  CCACATCATCAACAAGAAATGCCATTGTCAACTGTGCTGAGATGGTGACCTTCGACATCATCA 837
OY      683  aacactacttgatctatgatatatacaaccacttgagagacaatatatagtactcaacggttatcaa 742
Db      838  AGGAAGAAGTGTGAGAGTCTACACTGTTTACTGACAACTTCCCTGTCACCTTGTGTCG 897

```

QY 743 gttatctcttcgacatgtagcttcatctcttggaacaccacgcatgcatcaaaaga 802

Db 898 cctttggagctgccttctgtgcgcacacagtggtggccctcccccggatgtggttaaaagacc 957

QY 803 gaataatgaaatcaaccacagatataaacaagaaggaaggaactttgtatataatcatcagctg 862

Db 958 gatcattgaagacgctccctccatgcagcaggtaccgcgacgccccttgc----- 999

QY 863 acgtctgattcagcctgcttccaagtggaagatcatactgtatataaaagcttlttac 922

Db 1000 actctatgcctgaagatgacggcctcagagagagacacacggccctttacaaagattttgggc 1059

QY 923 catcttgctgagaaatgacccctctgttcaatggtgtctgcgttaactatgaaaaatca 982

Db 1060 cctctctttctgcgctcgtggagactttggaaagactatataattgttttaacatrttagcaactgca 1119

```

1      RESULT 13
2      -08-937-466-5
3      Sequence 5, Application US/08937466
4      Patent No. 5846779
5      GENERAL INFORMATION:
6      APPLICANT: Zhang, Ning
7      APPLICANT: Amaral, M. Catherine
8      APPLICANT: Chen, jin-long
9      TITLE OF INVENTION: UCP3 Genes
10     NUMBER OF SEQUENCES: 6
11     CORRESPONDENCE ADDRESS:
12     ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
13     STREET: 75 DENISE DRIVE
14     CITY: HILLSBOROUGH
15     STATE: CALIFORNIA
16     COUNTRY: USA
17     ZIP: 94010
18     COMPUTER READABLE FORM:
19     MEDIUM TYPE: Floppy disk
20     COMPUTER: IBM PC compatible
21     OPERATING SYSTEM: PC-DOS/MS-DOS
22     SOFTWARE: Patentin Release #1.0, Version #1.30
23     CURRENT APPLICATION DATA:
24     APPLICATION NUMBER: US/08/937,466
25     FILING DATE:
26     CLASSIFICATION: 435
27     ATTORNEY/AGENT INFORMATION:
28     NAME: OSMAN, RICHARD A
29     REGISTRATION NUMBER: 36,627
30     REFERENCE/DOCKET NUMBER: T97-009
31     TELECOMMUNICATION INFORMATION:
32     TELEPHONE: (650) 343-4341
33     TELEFAX: (650) 343-4342
34     INFORMATION FOR SEQ ID NO: 5:
35     SEQUENCE CHARACTERISTICS:
36     LENGTH: 1777 base pairs
37     TYPE: nucleic acid
38     STRANDEDNESS: double
39     TOPOLOGY: linear
40     MOLECULE TYPE: cDNA
41     -08-937-466-5

```

Query Match	4.7%	Score 48.6	DB 2	Length 1777
Best Local Similarity	48.1%	Pred. No. 1.4e-06		
Matches 138	Conservative	0	Mismatches 149	Indels 0
				Gaps 0

Dy	623	cgaatatacaaaagcgacgactgtgtatataatgtagagttaaccacttatgatcacgtga	682
QY	778	ccaaatcatcacaagaattgccatttcttcacactgtcttgagattggtaacctgcacacttcawca	837
Dy	718	atgccttacagaaacatcgcccacggaggagagatcagagccctttggamaagcacttggc	777
QY	563	atgcatttcgaaaactcttagctbaaagaggaataacagagagctttggcagcgctgggtac	622

QY 683 aaacactcttgatattaacacacactctgagacaaatcatgatgacacaggtttacaa 742

Db 838 AGGAAAGAACTCTGGAGTCTTACCTGTTTACTGCAACTTCCCTGTCACTTTGTCTCG 897

QY 743 gttatgtctctcgagctgtagctctctatctctggaacacacacgagatgtatcaaaagca 802

Db 898 CCTTTGGAGCTGGCTTCTGTGCCCACTGGTGGGCTTCCCGGTGGATGTGTAAAGACC 957

QY 803 gaataatgaatcaacacagagataaacaagaaagggagcttttgat 849

Db 958 GATACATGAAGCCTCCCTAAGCAGGTACCGCAGCAGAGACTCAGAAT 1004

```

RESULT 14
US-09-172-528-5
Sequence 5, Application US/09172528
Patent No. 5952469
GENERAL INFORMATION:
APPLICANT: Zhang, Ning
APPLICANT: Amaral, M. Catherine
APPLICANT: Chen, Jin-Long
TITLE OF INVENTION: ucp3 Genes
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/172,528
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/937,466
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEO ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1777 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-172-528-5

```

Query Match	4.7%	Score	48.6	DB 2	Length	1777
Best Local Similarity	48.1%	Pred.	NC.1.4e-06			
Matches 138; Conservative	0	Mismatches	149	Indels	0	Gaps 0

[illegible]

Db 838 AGGAGAAGTGTGCTGAGCTTCACCTGTTTACTGACAACTTCCCTGTGCTCTCTG 897  
 QY 743 gttatgtcttgagctgtagcttctatctctggaacaccgagatgcatcaaaaga 802  
 Db 898 CCTTTGAGCTGGCTCTGTGCTGACCACTGCTGCTCCCGGTGAGTGTGTAAGACCC 957  
 QY 803 gaataatgaatcaaccagagataaacaagaaggagacttggat 849  
 Db 958 GATACATGAACGCTCCCTAGGCAAGTACCGCACAGACTCAGAAAT 1004

RESULT 15

US-09-318-199-5  
 ; Sequence 5, Application US/09318199  
 ; Patent No. 6025469  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Zhang, Ning  
 ; APPLICANT: Amaral, M. Catherine  
 ; APPLICANT: Chen, Jin-Long  
 ; TITLE OF INVENTION: UCP3 Genes  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
 ; STREET: 75 DENISE DRIVE  
 ; CITY: HILLSBOROUGH  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94010  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/318,199  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/937,466  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OSMAN, RICHARD A  
 ; REGISTRATION NUMBER: 36,627  
 ; REFERENCE/DOCKET NUMBER: T97-009  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (650) 343-4341  
 ; TELEFAX: (650) 343-4342  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1777 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; US-09-318-199-5

Query Match 4.7%; Score 48.6; DB 3; Length 1777;

Best Local Similarity 48.1%; Pred. No. 1.4e-06;

Matches 138; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 563 atgcattgcacaaatcttgctgaagaggaataacgagggcttggcaggtcggtac 622  
 Db 718 ATGCTTACAGAACCTGAGGAGGAGAGAGTCAAGGGCTGTGGAAGGAGCTTGGC 777  
 QY 623 ccaatatacaaaagagcagctggtgaatatggagatttaaccactatgatacagta 682  
 Db 778 CCAACATCACAGAAATGCACTTGTCACTGTGCTGAGATGTGACCTACGACATCA 837  
 QY 683 aacactacttgatltgaataaccacttgaagacaatatcatcactacaggttataca 742  
 Db 838 AGGAGAAGTGTGCTGAGCTTCACCTGTTTACTGACAACTTCCCTGTGCTCTCTG 897

QY 743 gttatgtcttgagctgtagcttctatctctggaacaccgagatgcatcaaaaga 802  
 Db 898 CCTTTGAGCTGGCTCTGTGCTGACCACTGCTGCTCCCGGTGAGTGTGTAAGACCC 957  
 QY 803 gaataatgaatcaaccagagataaacaagaaggagacttggat 849  
 Db 958 GATACATGAACGCTCCCTAGGCAAGTACCGCACAGACTCAGAAAT 1004

Search completed: April 28, 2001, 18:18:31  
 Job time: 2730 sec



Trying 3106016892...Open

Welcome to STN International! Enter xxx  
LOGINID:ssppa1816mt  
PASSWORD  
TERMINAL (ENTER 1, 2, 3, OR 7):2

\*\*\*\*\* Welcome to STN International \*\*\*\*\*

NEWS 1 Web Page URLs for STN Seminar Schedule - N America  
NEWS 2 Dec 17 The CA LEXICON available in the CAPLUS and CA files  
NEWS 3 Feb 06 Engineering Information Encopass files have new names  
NEWS 4 Feb 16 TOXLINE no longer being updated  
NEWS 5 Apr 23 Search Devent WPINDEX by chemical structure  
NEWS 6 Apr 23 PRE-1987 REFERENCES NOW SEARCHABLE IN CAPLUS  
AND CA

NEWS EXPRESS April 18 CURRENT WINDOWS VERSION IS V6.0  
CURRENT MACINTOSH VERSION IS V5.0C (ENG) AND V5.0AB (JP).  
AND CURRENT DISCOVER FILE IS DATED 04/06  
NEWS HOURS STN Operating Hours Plus Help Desk Availability  
NEWS INTER General Internet Information  
NEWS LOGIN Welcome Banner and News Items  
NEWS PHONE Direct Dial and Telecommunication Network Access to STN  
NEWS WWW CAS World Wide Web Site (General Information)

Enter NEWS followed by the item number or name to see news on that specific topic.

All use of STN is subject to the provisions of the STN Customer agreement. Please note that this agreement limits use to scientific research. Use for software development or design or implementation of commercial gateways or other similar uses is prohibited and may result in loss of user privileges and other penalties.

\*\*\*\*\* STN Columbus \*\*\*\*\*

FILE 'HOME' ENTERED AT 12:39:45 ON 03 MAY 2001

=> file medicine cancerlit biosis embase scisearch

COST IN U.S. DOLLARS	ENTRY	SESSION	SINCE FILE	TOTAL
FULL ESTIMATED COST	0.15	0.15		

FILE 'MEDLINE' ENTERED AT 12:40:09 ON 03 MAY 2001

FILE 'CANCERLIT' ENTERED AT 12:40:09 ON 03 MAY 2001

FILE 'BIOSIS' ENTERED AT 12:40:09 ON 03 MAY 2001

COPYRIGHT (C) 2001 BIOSIS(R)

FILE 'EMBASE' ENTERED AT 12:40:09 ON 03 MAY 2001

COPYRIGHT (C) 2001 Elsevier Science B.V. All rights reserved.

FILE 'SCISEARCH' ENTERED AT 12:40:09 ON 03 MAY 2001

COPYRIGHT (C) 2001 Institute for Scientific Information (ISI) (R)

=> s ucpr4

L1 29 UCPR4

=> dup term

ENTER 'L# LIST OR (END):11

L2 10 DUP REM L1 (19 DUPLICATES REMOVED)

=> d 12 1-10 bib ab

L2 ANSWER 1 OF 10 MEDLINE DUPLICATE 1  
AN 2001227289 MEDLINE  
DN 21134422 PubMed ID: 11239468  
TI Homologues of the uncoupling protein from brown adipose tissue (UCP1):

UCP2, UCP3, BMCP1 and UCP4.

AU Bouilland F, Couplan E, Pecqueur C, Riquier D  
CS CEREMOD, C.N.R.S., UPR 9078, 9 rue Jules Hetzel, 82190 Meudon, France.  
bouilland@intbiogen.fr  
SO BIOCHIMICA ET BIOPHYSICA ACTA, (2001 Mar 1) 1504 (1) 107-119. Ref: 83  
Journal code: AOW; 0271513. ISSN: 0006-3002.

CY Netherlands

DT Journal; Article; (JOURNAL ARTICLE)

LA English

FS Priority Journals

EM 200104

EW Entered STN: 20010502

Last Updated on STN: 20010502

Entered PubMed: 20010312

Entered Medline: 20010426

L2 ANSWER 2 OF 10 MEDLINE DUPLICATE 2

AN 2000412260 MEDLINE

DN 20387222 PubMed ID: 10928996

TI Characterization of novel UCP5/BMCP1 isoforms and differential regulation of UCP4 and UCP5 expression through dietary or temperature manipulation.

AU Yu X X, Mao W, Zhong A, Schow P, Brush J, Sherwood S W, Adams S H, Pan G

CS Departments of Endocrinology, Molecular Biology, and Bioassay and Biomege, Genentech, Inc., South San Francisco, California 94080, USA.

SO FASEB JOURNAL, (2000 Aug) 14 (11) 1611-8.

Journal code: FAS; 8804484. ISSN: 0892-6638.

CY United States

DT Journal; Article; (JOURNAL ARTICLE)

LA English

FS Priority Journals

EM 200008

EW Entered STN: 20000907

Last Updated on STN: 20000907

Entered Medline: 20000825

AB Mitochondrial uncoupling proteins have been implicated in the maintenance of metabolic rate and adaptational thermoregulation. We recently reported the identification of a brain-specific mitochondrial uncoupling protein homologue, UCP4. Here we characterized another newly described member of the uncoupling protein family, termed UCP5 (also called BMCP1), UCP5 transcripts are present in multiple human and mouse tissues, with an especially high abundance in the brain and testis. Expression of UCP5 in mammalian cells reduces the mitochondrial membrane potential. Multiple isoforms of UCP5 were identified and exhibited tissue-specific distribution and different potency in reduction of membrane potential. Furthermore, the mRNA abundance of both UCP4 and UCP5 is modulated by nutritional status or temperature in a tissue-specific manner in mice. Brain UCP4 and UCP5 mRNA transcripts rose by 1.5- and 1.7-fold, respectively, and liver UCP5 expression increased by 1.8-fold in response to acute cold exposure. A high-fat diet increased UCP5 mRNA in liver by 1.6-fold selectively in the obesity-resistant AU but not in the obesity-prone C57BL/6J mouse strain. Liver UCP5 expression decreased significantly with a 24 h fast and was restored to the normal level after refeeding. In contrast, brain transcripts for both genes were not significantly altered by fasting or high-fat diet. These findings are consistent with the notion that UCP4 and UCP5 may be involved in tissue-specific thermoregulation and metabolic changes associated with nutritional status.

L2 ANSWER 3 OF 10 MEDLINE DUPLICATE 3

AN 2000202381 MEDLINE

DN 20202381 PubMed ID: 10736318

TI Uncoupling protein homologs: emerging views of physiological function.

AU Adams S H

CS Department of Endocrinology, Genentech, Incorporated, South San Francisco, CA 94080, USA.

SO JOURNAL OF NUTRITION, (2000 Apr) 130 (4) 711-4. Ref: 57  
Journal code: JEV; 0040243. ISSN: 0022-3166  
CY United States  
DT Journal; Article; (JOURNAL ARTICLE)

General Review; (REVIEW)

LA English

FS Priority Journals

EM 200004

EW Entered STN: 20000505

Last Updated on STN: 20000505

Entered Medline: 20000424

AB The widespread occurrence of excess weight and related diseases demands that efforts be made to understand energy expenditure from the gene to the whole animal. For some time, it has been understood that mitochondrial oxidation of fuels generates an electrochemical gradient via outward pumping of protons by the electron transport chain. ATP production via F(1)F(0) ATP synthase is then facilitated by the inward flux of protons down the gradient. There is a growing appreciation that a significant portion of the metabolic rate of endotherms is attributable to counteracting "proton leak" (uncoupling), wherein a flux of protons down the electrochemical gradient generates heat independently of ATP production. Proton leak is especially apparent in thermogenic brown adipose tissue, which expresses a tissue-specific uncoupling protein (UCP1). The recent discovery of widely expressed putative UCP1 homologs [UCP2, UCP3, UCP4, UCP5] brain mitochondrial carrier protein-1 (BMCP1)] raised the possibility that innate proton leak and metabolic rate are regulated by UCP1-like proteins. On the basis of current published data, one may not exclude the possibility that UCP homologs influence metabolic rate.

L2 ANSWER 4 OF 10 MEDLINE DUPLICATE 4

AN 2000233287 MEDLINE

DN 20233287 PubMed ID: 10772343

TI Specific sequence of motifs of mitochondrial uncoupling proteins.

AU Jezek P, Urbankova E

CS Department of Membrane Transport Biophysics, Institute of Physiology, Academy of Sciences of the Czech Republic, Prague.

SO JUBMB Life, (2000 Jan 49 (1) 63-70

Journal code: DJT; 10088706. ISSN: 1521-6543.

CY ENGLAND: United Kingdom

DT Journal; Article; (JOURNAL ARTICLE)

LA English

FS Priority Journals

EM 200006

EW Entered STN: 20000706

Last Updated on STN: 20000706

Entered Medline: 20000626

AB We have searched for the exclusivity of common sequence motifs of the mitochondrial uncoupling proteins (UCP1, UCP2, UCP3, UCP4, BMCP1, and plant UCP [PUMF]) within the gene family of mitochondrial anion carrier proteins. The UCP-specific sequences, "UCP signatures", were found in the first, second, and fourth alpha-helices. First: Ala/Ser-Cys/Thr/n-Phe-Ala/Gly-(negatively charged residue)-n-Phe-n-Cys-Thr-Phe/n, second: Gly/Ala-Ileu/n-Gln-X-(positively charged residue)-n-H-n-Cys-Ser/n-Phe/n, fourth: Pro-Asn/Thr/n-X-(positively charged residue)-Ileu/Met-Gly/Val-n-Thr, fourth: Pro-Asn/Thr/n-X-(negatively charged residue)-Asn/Ser/Ala-n-Ileu/n-Asn/Val-Cys/n-n-Thr-(negatively charged residue)-n-Thr/Pro-OH/Val (n, nonpolar, pH, aromatic; (positively charged residue/negatively charged residue, charged residue). The second and part of the third signature are also present in the yeast dicarboxylate transporter. The UCP signature excluding BMCP1 was also found in the second matrix segment, (positively charged residue)-Pro-Ileu/Leu/Ileu-(positively charged residue)-pH-X-Gly/Ser-Thr/n-X-NH-(negatively charged residue)-Ala-pH. These UCP signatures are thought to be involved in fatty acid anion binding and translocation.

L2 ANSWER 5 OF 10 BIOSIS COPYRIGHT 2001 BIOSIS

AN 2001 88714 BIOSIS

DN PREV200100088714

TI BMCP1: a neuronal protein which regulates mitochondrial free radical production and respiratory rate.

AU Kim-Han, J. S. (1); Reichert, S. A.; Dugan, L. L.

CS (1) Washington University School of Medicine, Saint Louis, MO USA  
SO Society for Neuroscience Abstracts, (2000) Vol. 26, No. 1-2, pp Abstract No. 378.16, print.

Meeting Info.: 30th Annual Meeting of the Society of Neuroscience New Orleans, LA, USA November 04-08, 2000 Society for Neuroscience



ISSN: 0190-5295.

DT Conference

LA English

SL English

AB Mitochondrial uncoupling proteins (UCPs) have been identified in various tissues, including brain. UCPs were first identified in brown adipose tissue, where they are thermogenic by uncoupling mitochondrial respiration from ATP production. Brain mitochondrial carrier protein 1 (BMCP1) is related to the UCPs, showing protein sequence homology of 34, 38, 39, and 39% with UCP1, UCP2, UCP3, and UCP4, respectively. We examined the function and characteristics of BMCP1 in brain using cortical neuronal cultures, and mouse and human brain tissue. We raised rabbit polyclonal antibody against the 14 amino acids corresponding to the protein sequence between the 1st and 2nd transmembrane domains of mouse BMCP1. This

AB

identified a single band of 36 kDa on western blot analysis of both adult mouse and human cortex, with lower levels of expression in mouse heart and kidney, and no expression in liver or lung. BMCP1 was expressed only in neurons, but not astrocytes, and protein expression in both cortical neurons and mouse brain was developmentally upregulated. BMCP1 was localized to mitochondria, as shown by immunocytochemistry and western blot analysis. Overexpression of BMCP1 in GT1-1 cells resulted in a significant reduction of basal superoxide production determined by dihydroethidium oxidation. Oxygen consumption rates using succinate as substrate were not affected by overexpression of BMCP1. However, oxygen consumption by BSA was more profound in BMCP1 cells than in control cells. Because chronic overexpression of BMCP1 appeared to affect the phenotype and proliferation of GT1-1 cells, we began to study the effect of induced BMCP1 overexpression using tetracycline-inducible PC12 transfectants.

L2 ANSWER 6 OF 10 MEDLINE

AN 2000117196 MEDLINE

DN 20117196 PubMed ID: 1063471

TI Uncoupling protein-a useful energy dissipator.

AU Klingenberg M

CS Institute of Physical Biochemistry, University of Munich, Germany.

SO JOURNAL OF BIOENERGETICS AND BIOMEMBRANES. (1998 Oct) 31

(5) 419-30. Ref:

80

Journal code: HIC; 7701859. ISSN: 0145-479X.

CY United States

DT Journal; Article. (JOURNAL ARTICLE)

General Review; (REVIEW)

(REVIEW, TUTORIAL)

LA English

FS Priority Journals

EM 200003

ED Entered STN: 20000320

Last Updated on STN: 20000320

Entered Medline: 20000306

AB The structure/function relationship in the uncoupling proteins (UCP) is reviewed, stressing UCP from brown adipose tissue (UCP1) since, so far, nearly no biochemistry is known for the UCP variants UCP2, UCP3, and UCP4. The transport for H<sup>+</sup> and Cl<sup>-</sup> and its dependence on fatty acids in reconstituted vesicles is described. The inhibition and binding of nucleotides to UCP1, in particular, the pH dependence and two-stage binding are analyzed. A model for the role of fatty acid in H<sup>+</sup> transport is shown. The role of specific residues in UCP1 is analyzed by directed mutagenesis in a yeast expression system. The different regulation by the cellular energy potential of UCP1 versus UCP3 is discussed.

L2 ANSWER 7 OF 10 MEDLINE

AN 1989148824 MEDLINE

DN 99148824 PubMed ID: 10025957

TI UCP4, a novel brain-specific mitochondrial protein that reduces

membrane potential in mammalian cells.

CM Erratum in: FEBS Lett 1989 Apr 23;449(2-3):283

AU Mao W, Yu X, X, Zhong A, Li W, Brush J, Sherwood S W, Adams S H, Pan

G

CS Department of Endocrinology, Genentech, Inc., South San Francisco, CA

94080, USA.

SO FEBS LETTERS. (1989 Jan 28) 443 (3) 326-30.

Journal code: EUH; 0155157. ISSN: 0014-5793.

CY Netherlands

DT Journal; Article. (JOURNAL ARTICLE)

FS Priority Journals

OS GENBANK:AF110532

EM 199903

ED Entered STN: 19990324

Last Updated on STN: 20000303

Entered Medline: 19990311

AB Uncoupling proteins (UCPs) are a family of mitochondrial transporter proteins that have been implicated in thermoregulatory heat production and maintenance of the basal metabolic rate. We have identified and partially characterized a novel member of the human uncoupling protein family, termed uncoupling protein-4 (UCP4). Protein sequence analyses showed that UCP4 is most related to UCP3 and possesses features characteristic of mitochondrial transporter proteins. Unlike other known UCPs, UCP4 transcripts are exclusively expressed in both fetal and adult brain tissues. UCP4 maps to human chromosome 9p11.2-q12. Consistent with its potential role as an uncoupling protein, UCP4 is localized to the mitochondria and its ectopic expression in mammalian cells reduces mitochondrial membrane potential. These findings suggest that UCP4 may be involved in thermoregulatory heat production and metabolism in the brain.

L2 ANSWER 8 OF 10 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.

AN 1999146875 EMBASE

TI Erratum: UCP4, a novel brain-specific mitochondrial protein that

reduces membrane potential in mammalian cells (FEBS Letters 443 (1999)

(326-330) PII: S001457939801713X)

AU Mao W, Yu X, X, Zhong A, Li W, Brush J, Sherwood S W, Adams S H, Pan

G

CS G. Pan, Department of Endocrinology, Genentech, Inc., 1 DNA Way, South

San

Francisco, CA 94080, United States, jippan@gene.com

SO FEBS Letters. (1999) 449/2-3 (293).

Ref: 0

ISSN: 0014-5793. CODEN: FEELBL

PUI S 0014-5793(99)00318-X

CY Netherlands

DT Journal; Errata

FS 029 Clinical Biochemistry

LA English

L2 ANSWER 9 OF 10 SCISEARCH COPYRIGHT 2001 ISI (R)

AN 1999-360327 SCISEARCH

GA The Genuine Article (R) Number: 192LW

TI UCP4, a novel brain-specific mitochondrial protein that reduces

membrane potential in mammalian cells (vol 443, pg 326, 1999)

AU Mao W G, Yu X, X, Zhong A, Li W, Brush J, Sherwood S W, Adams S H, Pan G

H (Reprint)

CS GENENTECH INC, DEPT ENDOCRINOL, M-S-371, 1 DNA WAY, S SAN

FRANCISCO, CA

94080 (Reprint); GENENTECH INC, DEPT ENDOCRINOL, S SAN

FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA





General Review, (REVIEW)  
(REVIEW, TUTORIAL)

LA English  
FS Priority Journals  
EM 200104  
ED Entered STN: 20010502  
Last Updated on STN: 20010502  
Entered PubMed 20010312  
Entered Medline: 20010426

L4 ANSWER 3 OF 8 MEDLINE  
AN 2000412260 MEDLINE  
DN 20387222 PubMed ID: 10928996  
TI Characterization of novel UCP5/BMCP1 isoforms and differential regulation of UCP4 and UCP5 expression through dietary or temperature manipulation, AU Yu X; Mao W; Zhong A; Schow P; Bush J; Sherwood S W; Adams S H; Pan G  
CS Departments of Endocrinology, Molecular Biology, and Bioassay and Biomeq, Genentech, Inc., South San Francisco, California 94080, USA.  
SO FASEB JOURNAL, (2000 Aug) 14 (11) 1611-8.  
CY United States  
DT Journal Article, (JOURNAL ARTICLE)  
English  
Priority Journals  
200008  
ED Entered STN: 20009907  
Last Updated on STN: 20009907  
Entered Medline: 20000825

AB Mitochondrial uncoupling proteins have been implicated in the maintenance of metabolic rate and adaptational thermoregulation. We recently reported the identification of a brain-specific mitochondrial uncoupling protein homologue, UCP4, here we characterized another newly described member

of the uncoupling protein family, termed UCP5 (also called BMCP1). UCP5 transcripts are present in multiple human and mouse tissues, with an especially high abundance in the brain and testis. Expression of UCP5 in mammalian cells reduces the mitochondrial membrane potential. Multiple isoforms of UCP5 were identified and exhibited tissue-specific distribution and different potency in reduction of membrane potential. Furthermore, the mRNA abundance of both UCP4 and UCP5 is modulated by nutritional status or temperature in a tissue-specific manner in mice.

Brain UCP4 and UCP5 mRNA transcripts rose by 1.5- and 1.7-fold, respectively, and liver UCP5 expression increased by 1.8-fold in response to acute cold exposure. A high-fat diet increased UCP5 mRNA in liver by 1.6-fold selectively in the obesity-resistant AU but not in the obesity-prone C57BL/6J mouse strain. Liver UCP5 expression decreased significantly with a 24 h fast and was restored to the normal level after refeeding. In contrast, brain transcripts for both genes were not significantly altered by fasting or high-fat diet. These findings are consistent with the notion that UCP4 and UCP5 may be involved in tissue-specific thermoregulation and metabolic changes associated with nutritional status.

L4 ANSWER 4 OF 8 BIOSIS COPYRIGHT 2001 BIOSIS

AN 2000197467 BIOSIS  
DN PRE200000187467  
TI Uncoupling protein homologs: Emerging views of physiological function, AU Adams, Sean H, (1)  
CS (1) Department of Endocrinology, Genentech, Incorporated, South San Francisco, CA, 94080 USA  
SO Journal of Nutrition, (April, 2000) Vol. 130, No. 4, pp. 711-714.  
ISSN: 0022-3166.

DT Article  
LA English  
SL English  
AB The widespread occurrence of excess weight and related diseases demands that efforts be made to understand energy expenditure from the gene to the whole animal. For some time, it has been understood that mitochondrial oxidation of fuels generates an electrochemical gradient via outward pumping of protons by the electron transport chain. ATP production via F1F0 ATP synthase is then facilitated by the inward flux of protons down the gradient. There is a growing appreciation that a significant portion of the metabolic rate of endotherms is attributable to countercoupling

"proton leak" (uncoupling), wherein a flux of protons down the electrochemical gradient generates heat independently of ATP production. Proton leak is especially apparent in thermogenic brown adipose tissue, which expresses a tissue-specific uncoupling protein (UCP1). The recent discovery of widely expressed putative UCP1 homologs (UCP2, UCP3, UCP4, UCP5/brain mitochondrial carrier protein-1 (BMCP1)) raised the possibility that innate proton leak and metabolic rate are regulated by UCP1-like proteins. On the basis of current published data, one may not exclude the possibility that UCP homologs influence metabolic rate.

L4 ANSWER 5 OF 8 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.  
AN 2000382329 EMBASE  
TI Appetite regulation and uncoupling proteins, AU ERLINGER OCH UKROPPLANDE PROTEINER, AU Erlanson-Albertsson C.  
CS Dr. C. Erlanson-Albertsson, Inst. for Cell och Molekylarbiologi, Lunds Universitet, P.O. Box 94, S-221 00 Lund, Sweden. charlotte.erlanson-albertsson@medkem.lu.se  
SO Scandinavian Journal of Nutrition/Hairingsforskning, (2000) 44(3) (108-110), Refs: 31  
ISSN: 1102-6480 CODEN: SJNUIE  
CY Sweden  
DT Journal: General Review  
FS 002 Physiology  
U29 Clinical Biochemistry  
048 Gastroenterology  
LA Swedish  
SL English  
AB Appetite regulation and feeding behavior is critical for survival. In understanding the mechanism for appetite regulation various theories have been postulated. One of these, the thermodynamic theory, postulates that appetite regulation is tightly linked to body temperature, a rise in body temperature causing the termination of a meal. With the discovery of a family of uncoupling proteins, UCP1-UCP5, the production of heat during feeding has attracted new interest. Special interest has been focused on UCP2, which is produced in the gastrointestinal tract.

L4 ANSWER 6 OF 8 BIOSIS COPYRIGHT 2001 BIOSIS  
AN 2000253107 BIOSIS  
DN PRE200000253107  
TI Specific sequence motifs of mitochondrial uncoupling proteins, AU Jezek, Petr (1); Urbanikova, Eva  
CS (1) Department of Membrane Transport Biophysics, Institute of Physiology, Academy of Sciences of the Czech Republic, Videnstva 1083, CZ 14220, Prague  
Czech Republic  
SO IUBMB Life, (Jan., 2000) Vol. 49, No. 1, pp. 63-70, print.  
ISSN: 1521-6543.

DT Article  
LA English  
SL English  
AB We have searched for the exclusivity of common sequence motifs of the mitochondrial uncoupling proteins (UCP1, UCP2, UCP3, UCP4, BMCP1, and plant UCP (PUMPT)) within the gene family of mitochondrial anion carrier proteins. The UCP-specific sequences, "UCP signatures", were found in the first, second, and fourth alpha-helices. First: Ala/Ser-Cys/Thr/n-Phe-  
AlaGly(-)-n-Phe-n-Cys-Thr-Phe/n, second: Gly/Ala-Ileu-Glu-X(-)-NH-  
n-Cys-Ser/n-phi-X-n-Ser-OH/Gly-n-(+)-Ileu/Met-Gly/Val-n-Tyr; fourth:  
n-Pro-Asn/Thr-n-X-(+)-Asn/Ser/Ileu-n-Ileu-leu-n-Asn/Val-Cys/n-n-Tyr(-)-  
n-Tyr-Ileu-CH/Val (n, nonpolar; phi, aromatic; +, charged residue). The second and part of the third signature are also present in the yeast dicarboxylate transporter. The UCP signature excluding BMCP1 was also found in the second matrix segment (+)-Pro/Ileu-Ileu/Ideu(-)-(+)-phi-X-  
Gly/Ser-Thr/n-X-NH(-)-Ala-phi. These UCP signatures are thought to be involved in fatty acid anion binding and translocation.

L4 ANSWER 7 OF 8 MEDLINE  
AN 2000117196 MEDLINE  
DN 20117196 PubMed ID: 10653471  
TI Uncoupling protein--a useful energy dissipator, AU Klingenberg M  
CS Institute of Physical Biochemistry, University of Munich, Germany  
SO JOURNAL OF BIOENERGETICS AND BIOMEMBRANES, (1999 Oct) 31 (6) 419-30. Ref: 80  
Journal code: HIO, 7701859, ISSN: 0145-479X

CY United States  
DT Journal Article, (JOURNAL ARTICLE)  
General Review, (REVIEW)  
(REVIEW, TUTORIAL)  
LA English  
FS Priority Journals  
EM 200003  
ED Entered STN: 20000320  
Last Updated on STN: 20000320  
Entered Medline: 20000306

AB The structure/function relationship in the uncoupling proteins (UCP) is reviewed, stressing UCP from brown adipose tissue (UCP1) since, so far, nearly no biochemistry is known for the UCP variants UCP2, UCP3, and UCP4.  
The transport for H+ and Cl- and its dependence on fatty acids in reconstituted vesicles is described. The inhibition and binding of nucleotides to UCP1, in particular, the pH dependence and two-stage binding are analyzed. A model for the role of fatty acid in H+ transport is shown. The role of specific residues in UCP1 is analyzed by directed mutagenesis in a yeast expression system. The different regulation by the cellular energy potential of UCP1 versus UCP3 is discussed.

L4 ANSWER 8 OF 8 MEDLINE  
AN 1999148824 MEDLINE  
DN 99148824 PubMed ID: 10025957  
TI UCP4, a novel brain-specific mitochondrial protein that reduces membrane potential in mammalian cells, AU Mao W; Yu X; Zhong A; Li W; Brush J; Sherwood S W; Adams S H; Pan G  
CS Department of Endocrinology, Genentech, Inc., South San Francisco, CA 94080, USA.  
SO FEBS LETTERS, (1999 Jan 29) 443 (3) 326-30  
Journal code: EUH, 0156157, ISSN: 0014-5793.  
CY Netherlands  
DT Journal Article, (JOURNAL ARTICLE)  
LA English  
FS Priority Journals  
OS GENBANK-AF110532  
EM 199903  
ED Entered STN: 19990324  
Last Updated on STN: 20000303  
Entered Medline: 19990311

AB Uncoupling proteins (UCPs) are a family of mitochondrial transporter proteins that have been implicated in thermoregulatory heat production and maintenance of the basal metabolic rate. We have identified and partially characterized a novel member of the human uncoupling protein family, termed uncoupling protein-4 (UCP4). Protein sequence analyses showed that UCP4 is most related to UCP3 and possesses

features characteristic of mitochondrial transporter proteins. Unlike other known UCPs, UCP4 transcripts are exclusively expressed in both fetal and adult brain tissues. UCP4 maps to human chromosome 6p11.2-q12. Consistent with its potential role as an uncoupling protein, UCP4 is localized to the mitochondria and its ecopic expression in mammalian cells reduces mitochondrial membrane potential. These findings suggest that UCP4 may be involved in thermoregulatory heat production and metabolism in the brain.

=> logoff  
ALL L# QUERIES AND ANSWER SETS ARE DELETED AT LOGOFF  
LOGOFF? (Y/N)HOLD Y  
COST IN U.S. DOLLARS ENTRY SESSION TOTAL  
FULL ESTIMATED COST 23.49 23.64  
STN INTERNATIONAL LOGOFF AT 12:41:31 ON 03 MAY 2001

